Supplementary Information

Dynamic covalent self-assembly of mechanically interlocked

molecules solely made from peptides

Hendrik V. Schröder[†], Yi Zhang[†], A. James Link[†,‡,§,*]

[†]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544,

United States

[‡]Department of Chemistry, Princeton University, Princeton, NJ 08544, United States

[§]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, United States

*Corresponding author: ajlink@princeton.edu

Table of Contents

1. Additional Methods

General Information

For DNA extraction from cells and DNA purification from agarose gels, DNA pruification kits from QIAGEN (Hilden, Germany) and Zymo Research (CA, USA) were used, respectively. The restriction enzymes *Xho*I and *Hin*dIII, Q5 high-fidelity DNA polymerase, and T4 DNA ligase were obtained from New England Biolabs (MA, USA). Molecular cloning was carried out with XL1-Blue *Escherichia coli*. Lasso peptide variants were expressed BL21 *E. coli*. Sanger sequencing was performed by Genewiz (NJ, USA). For HPLC and LC-MS, HPLC grade solvents were used. Acetonitrile (ACN) was purchased from Sigma Aldrich (MO, USA). 1,4-Dithiothreitol (DTT) and iodoacetamide was obtained from Oakland Chemicals (MI, USA). Proteolytic enzymes trypsin and thermolysin were purchased from Promega (WN, USA) and Sigma Aldrich (MO, USA), respectively. Carboxypeptidase was purchased from Affymetrix (CA, USA).

HPLC

Lasso peptide purification was performed on Agilent 1200 HPLC system using a Zorbax SB-300-C18 semi-prep HPLC Column (9.4 × 250 mm, 5 μm particle size) with a binary gradient consisting of water and ACN both containing 0.1 % (v/v) trifluoroacetic acid. The linear, binary solvent gradient (flow rate 4 mL/min) was applied as follows: From 0–1 min, the solvent was maintained as 90% water and 10% ACN. This was followed by linear ramping to 50% water and 50% ACN from 1–20 min and subsequent ramping to 10% water and 90% ACN (20–25 min). Afterwards, the solvent compostion was kept constant (25–30 min) and gradually changed again to 90% water and 10% ACN (30–32 min). Analytical runs were performed on a ZORBAX 300SB-C18 (3.0 × 150 mm, 3.5 μm) column using a decreased flow rate of 0.75 mL/min and the same solvent gradient.

LC/MS

For LC/MS measurements, reaction solutions were injected onto a Zorbax 300SB-C18 column (2.1 × 50 mm, 3.5 μm) using a Agilent 1260 system. The mobile phase consisted of a mixture of water and ACN (HPLC grade) both containing 0.1 % (v/v) formic acid. A linear, binary solvent gradient (flow rate 0.5 mL/min) was applied as follows. From 0–1 min, the solvent was maintained as 90% water and 10% ACN. This was followed by linear ramping to 50% water and 50% ACN from 1–20 min and subsequent ramping to 10% water and 90% ACN (20–25 min). Afterwards, the solvent compostion was kept constant for the rest of the run time (25–30 min). The separated species were analyzed with an in-line Agilent 6530 Accurate Mass Q-TOF mass

spectrometer using electrospray ionization in positive mode. MS[2] and CID experiments were performed with a manual ion selection and different collision voltages. Peptide concentrations were determined by UV/Vis absorbance at 280 nm with a Nanodrop spectrophotometer. Samples were lyophilized using a Labconco FreeZone freeze dry system.

Plasmid Construction

Plasmid construction was performed as described in the main text Methods and previously.[1] All MccJ25 mutants were created using primers found in Table 2 to amplify the MccJ25 precursor gene *mcjA*. This amplicon is flanked by 5' *Xho*I and 3' *Hin*dIII sites. Insert and vector pJP73 were digested by these enzymes and ligated together. All variants of MccJ25 were generated in a similar fashion (see Table 1). For variants with G21C substitution, an intermediate plasmid pHS09 was generated and used as template. Plasmid pCA45 was used as template for variants with G12R I13C substitutions. Sequences of all constructs were confirmed by Sanger sequencing (GENEWIZ, NJ, USA).

Kinetic Measurements

Digestion reactions, typically 50 µL, were placed into a sealed vial to prevent rapid oxidation in air. Afterwards, aliquots of the solution were anaylzed by HPLC at 1, 36, 72, 108, 144, 180, 216, 252, 288, 324, 360, 1440, and 5760 minutes. Furthermore, we investigated the reaction at early time points: 1, 5, 10, 15, and 30 minutes. For this, the trypsin digest and disulfide exchange was "frozen" by lowering the pH < 4 by addition of TFA (10 % v/v). Fractions of library members were determined by integration of their HPLC peak area.

NMR

After HPLC isolation, fractions were lyophilized and the resulting white powder was taken up in D_2O/H_2O (5/95, 170 µL) yielding a 1.03 mM aqueous solution as determined by optical absorbance at 280 nm and HPLC. Purity of the sample was tested before and after NMR experiments by LC/MS with no signs of degradation or impurities (Figure 10). 1D and 2D (TOCSY, NOESY, EXSY) NMR spectra were recorded on an 800 MHz Bruker Ascend NMR spectrometer. The water peak was suppressed in all spectra. The NOESY and EXSY spectra were acquired using mixing times (t_m) of 300 and 1500 ms, respectively, at 295 K. Variable temperature (VT) NMR experiments were performed in a temperature range of 277 to 337 K. Subsequent cooling to room temperature re-generated the initial ^{1}H NMR spectrum.

Solution NMR Structure Calculation of [2]Catenane 3-2H

Chemical shifts for protons of both conformers were separately assigned on the basis of intra- (TOCSY) and inter-residue (NOESY) correlations. Due to partial peak overlapping, some peaks could not be assigned with full certainty (91 % peak assignment for the major conformer). Off-diagonal crosspeaks, corresponding to NOE interactions, were manually picked and integrated from the NOESY spectrum. These peak volumes were then used for distance calibration and as upper distance restraints for structural calculations using the CYANA software package.[2] Seven cycles of automated NOE peak assignment and structure calculations were performed starting with 100 initial structures, followed by a final structural calculation. The top 20 structures were subsequently energy-minimized in explicit solvent according to a procedure by Spronk *et al.*[3,4] Briefly, the [2]catenane was placed in a solvent box with a 10 Å layer of TIP3P water in each direction of the peptide. The system was autoionized by adding NaCl. Using the CHARMM force field, the NAMD[5] (version 2.13) and VMD[6] (version 1.93) software packages, an annealing procedure followed by a final minimization step was applied. The isopeptide-bond (Gly-1 to Glu-8) was force field parametrized according to the reported procedure for MccJ25 of Lai *et al.*[7]

MD Simulations

The CHARMM force field and the NAMD[5] (version 2.13) software package were used for MD simulations. First, structures were energy minimized in 10000 steps followed by a 29800-step heating procedure, ramping from 0 to 298 K. Subsequently, a 1 ns equilibration run was performed at 298 K. The following 1 μs production run was performed in the NPT ensemble to reproduce the experimental conditions. Pressure and temperature control was achieved by using a Langevin piston (1 atm, 100 fs piston period, 50 fs piston decay) and Langevin thermostat (298 K, damping coefficient of 1 ps^{-1}). Full electrostatic interactions were taken into account using the Particle Mesh Ewald (PME) algorithm with a grid spacing of 1 Å. Force-field parameters of the simulation were set to exclude = scaled 1-4, a 1-4 scaling value of 0.5, a van der Waals interaction cut-off distance of 12 Å, a switch distance of 10.5 Å and pairlistdist value of 13.5 Å. The RATTLE algorithm was applied to constrain all bonds involving hydrogen atoms. Initially, five 10-ns-simulations were performed using the first five structures of the solution NMR ensemble as starting conformations for the MD simulations in water. A structural comparison of the simulation trajectories shows that all systems share a similar conformational space indicating that all of these structures are suitable starting conformations. Afterwards, the lowest-energy structure of the NMR ensemble was selected and subjected to the 1 μs production run. The resulting MD trajectories were analyzed using the VMD[6] (version 1.93) software package.

Briefly, the average RMSD value for the backbone of each residue was calculated over the full simulation time of 1 µs. To exclude co-conformational contributions on the B-factor mapping (Figure 46), RMSD values were referenced to the alignment of the backbone of indiviual peptide chains (isopeptide-bonded ring, residues 1–10; disulfide-bonded ring, residues 11–21).

2. Plasmids and Primers

Table 1. Plasmids used in this study

Plasmid	Template	Substitutions	Primers	Restriction sites	Vector
pHS09	pJP73	G21C	1, 5	*Xho*I / *Hin*dIII	pQE-80
pHS11	pHS09	H5C, G21C	1, 6, 9, 10	*Xho*I / *Hin*dIII	pQE-80
pHS10	pHS09	G12C, G21C	1, 6, 7, 8	*Xho*I / *Hin*dIII	pQE-80
pCA45[a]	/	G12R, I13C	/	*Xho*I / *Hin*dIII	pQE-60
pHS08	pCA45	H5C, G12R, I13C	1, 2, 3, 4	*Xho*I / *Hin*dIII	pQE-80
pHS06	pCA45	G12C, I13C, G21C	1, 5, 11, 12	*Xho*I / *Hin*dIII	pQE-80
pHS07	pCA45	H5C, G12R, I13C, G21C	1, 5, 9, 10	*Xho*I / *Hin*dIII	pQE-80
pHS15	pHS10	G12C, S18D, G21C	1, 13	*Xho*I / *Hin*dIII	pQE-80
pHS16	pHS10	G12C, P16A, G21C	1, 14	*Xho*I / *Hin*dIII	pQE-80

[a] Cloning of this plasmid has been previously described.[1]

Table 2. Primers used in this study

Primer	Sequence (5' to 3')
1	GGCGTATCACGAGGCCCTTTC
2	GCTAATTAAGCTTCAGCCATAGAAAGATATAGGTGTA
3	GGTGCAGGATGCGTGCCTGAGTATTTTG
4	CAAAATACTCAGGCACGCATCCTGCACC
5	CAGCTAATTAAGCTTCAGCAATAGAAAGATATAGGTGTACCGC
6	CGCCAAGCTAGCTTGGATTCTCACC
7	GTGCCTGAGTATTTTGTGTGCATTGG
8	CCAATGCACACAAAATACTCAGGCAC
9	GGTGCAGGATGCGTGCCTGAGTATTTTG
10	CAAAATACTCAGGCACGCATCCTGCACC
11	GAGTATTTTGTGTGCATTGG
12	CCAATGCACACAAAATACTC
13	ATTAAGCTTCAGCAATAGAAATCTATAGGTGTACCAATGC
14	ATTAAGCTTCAGCAATAGAAAGATATGGCTGTACCAATGC

Table 3. Sequencing data

Plasmid	Substitutions	Sequence
pHS09	G21C	NNNNNNNNNNNNCNTAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGAT TCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACCA TGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATCT CCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGGT GCAGGACATGTGCCTGAGTATTTTGTGGGGATTGGTACACCTATATCTTTCTATTGCT GAAGCTT
pHS11	H5C, G21C	NNNNNNNNNNNANNNNNNNNNTTATTTGCTTTGTGAGCGGATAACAATTATAATAGA TTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACC ATGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATC TCCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGG TGCAGGATGCGTGCCTGAGTATTTTGTGGGGATTGGTACACCTATATCTTTCTATTG CTGAAGCTT
pHS10	G12C, G21C	NNNNNNNNNNNNNNNNNNNNTTATTTGCTTTGTGAGCGGATAACAATTATAATAGA TTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACC ATGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATC TCCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGG TGCAGGACATGTGCCTGAGTATTTTGTGTGCATTGGTACACCTATATCTTTCTATTGC TGAAGCTT
pHS08	H5C, G12R, I13C	NNNNNNNNNNNNNNTNNNNNNNTTATTTGCTTTGTGAGCGGATAACAATTATAATAGA TTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACC ATGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATC TCCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGG TGCAGGATGCGTGCCTGAGTATTTTGTGCGCTGCGGTACACCTATATCTTTCTATGG CTGAAGCTT
pHS06	G12C, I13C, G21C	NNNNNNNNNNNNANNNNANTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA ATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACCATGA TTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATCTCCT GCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGGTGCA GGACATGTGCCTGAGTATTTTGTGTGCTGCGGTACACCTATATCTTTCTATTGCTGAA GCTT
pHS07	H5C, G12R, I13C, G21C	NNNNNNNNNNNNCNTNNNNNNNTTTATTTNCTTTGTGAGCGGATAACAATTATAATAG ATTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAAC CATGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCAT CTCCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTG GTGCAGGATGCGTGCCTGAGTATTTTGTGCGCTGCGGTACACCTATATCTTTCTATT GCTGAAGCTT
pHS15	G12C, S18D, G21C	NNNNNNNNNCNGATAACATTATAATAGATTCAATTGTGAGCGGATAACAATTTCACACA GAATTCATTAAAGAGGAGAAATTAACCATGATTAAGCATTTTCATTTTAATAAACTGTC TTCTGGTAAAAAAAATAATGTTCCATCTCCTGCAAAGGGGGTTATACAAATAAAAAAA TCAGCATCGCAACTCACAAAAGGTGGTGCAGGACATGTGCCTGAGTATTTTGTGTGC ATTGGTACACCTATAGATTTCTATTGCTGAAGCTT
pHS16	G12C, P16A, G21C	NNNNNNNNNNNCNTNNNNNANTTATTTGCTTTGTGAGCGGATAACAATTATAATAGA TTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACC ATGATTAAGCATTTTCATTTTAATAAACTGTCTTCTGGTAAAAAAAATAATGTTCCATC TCCTGCAAAGGGGGTTATACAAATAAAAAAATCAGCATCGCAACTCACAAAAGGTGG TGCAGGACATGTGCCTGAGTATTTTGTGTGCATTGGTACAGCCATATCTTTCTATTCT GAAGCTT

3. Purification and Characterization of MccJ25 Variants

Table 4. Expression levels and retention times of MccJ25 variants obtained by HPLC

Species	Substitution	Yield[a] / mg·L^{-1}	Retention time[b] / min
L2	H5C, G21C	0.53	18.2
L3	G12C, G21C	0.71	17.1
L4	G12R, I13C	0.73	14.4
L5	H5C, G12R, I13C	0.65	16.6
L6	G12C, I13C, G21C	0.34[c]	16.2
L7	H5C, G12R, I13C, G21C	0.62	17.3
L8	G12C, S18D, G21C	0.30	17.2

[a] Yields are reported as isolated yields and were determined after HPLC purification by UV/Vis measurements. The optical absorptions of the peptides were calculated according to their Tyr and Cys content. Yields were determined on basis of an expression with 0.5 L of culture.

[b] Retention times were determined by analytical HPLC as described above.

[c] Lasso peptide **L6** was HPLC-purified twice to remove coeluting impurities.

Table 5. High resolution mass spectrometrical data of variants **L2–8** and **L3 P16A**.

Species	Chem. formula	Observed m/z for $[M+2H]^{2+}$ and $[M+3H]^{3+}$	Calcd. m/z for $[M+2H]^{2+}$ and $[M+3H]^{3+}$	Δ / ppm
L2	$C_{99} H_{139} N_{21} O_{27} S_2$	1059.9872, 706.9940	1059.9868, 706.9936	0.4, 0.6
L3	$C_{103} H_{143} N_{23} O_{27} S_2$	1100.0059, 733.6731	1100.0055, 733.6728	0.4, 0.4
L4	$C_{102} H_{142} N_{26} O_{27} S$	1098.5208, 732.6830	1098.5202, 732.6826	0.5, 0.5
L5	$C_{99} H_{140} N_{24} O_{27} S_2$	1081.4976, 721.3359	1081.4953, 721.3326	2.1, 4.6
L6	$C_{100} H_{137} N_{23} O_{27} S_3$	1094.9684, 730.3173	1094.9681, 730.3145	0.3, 3.8
L7	$C_{100} H_{142} N_{24} O_{27} S_3$	1104.4887, 736.6632	1104.4892, 736.6619	0.5, 1.8
L8	$C_{104} H_{143} N_{23} O_{28} S_2$	1114.0036, 743.0048	1114.0030, 743.0044	0.5, 0.5
L3 P16A	$C_{101} H_{141} N_{23} O_{27} S_2$	1086.9984, 725.0015	1086.9977, 725.0009	0.6, 0.8
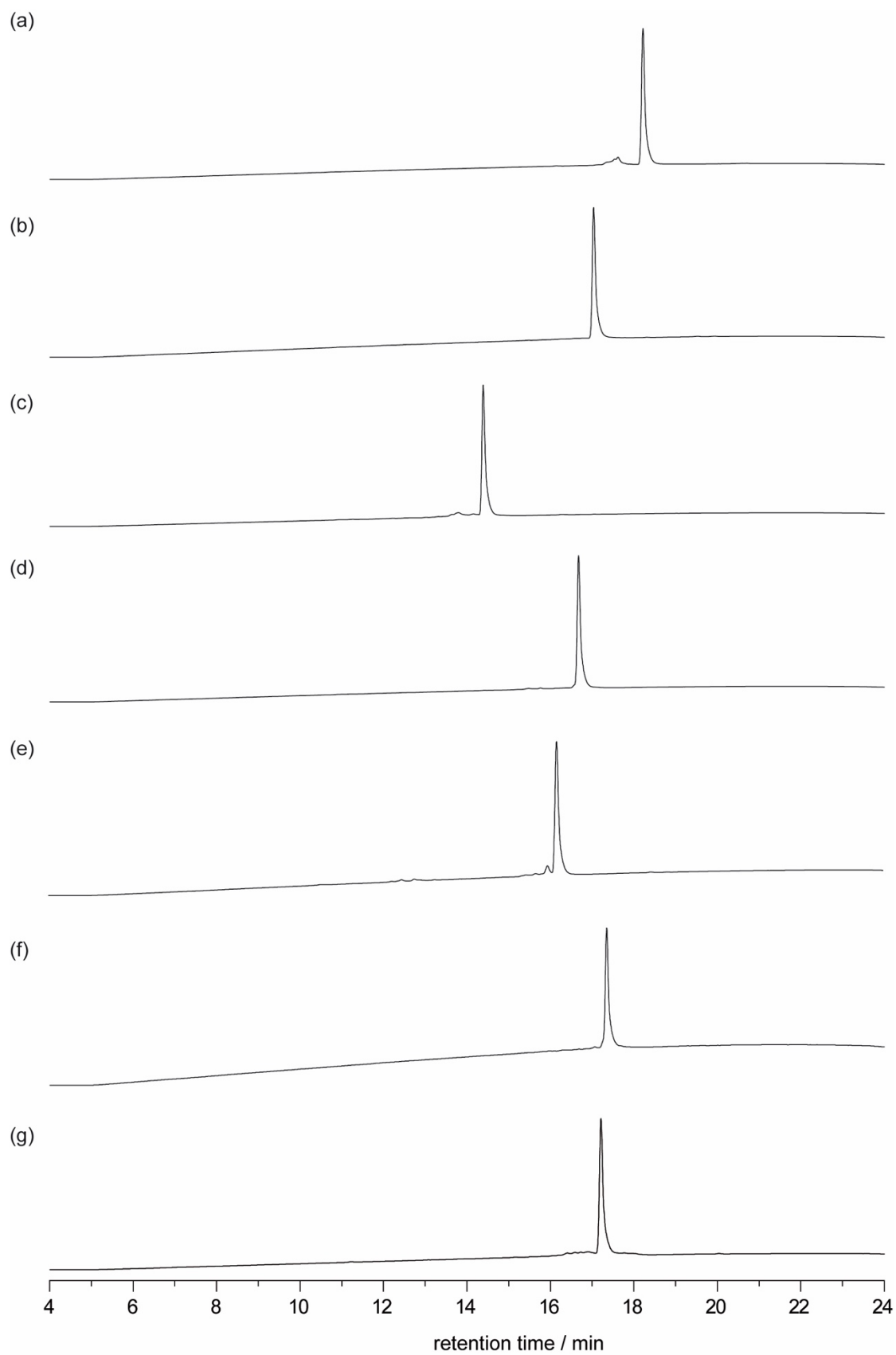
Figure 1. RP-HPLC traces of lasso peptide variants **L2**–**8** (a-g) used for DCL formation.

4. Additional Figures
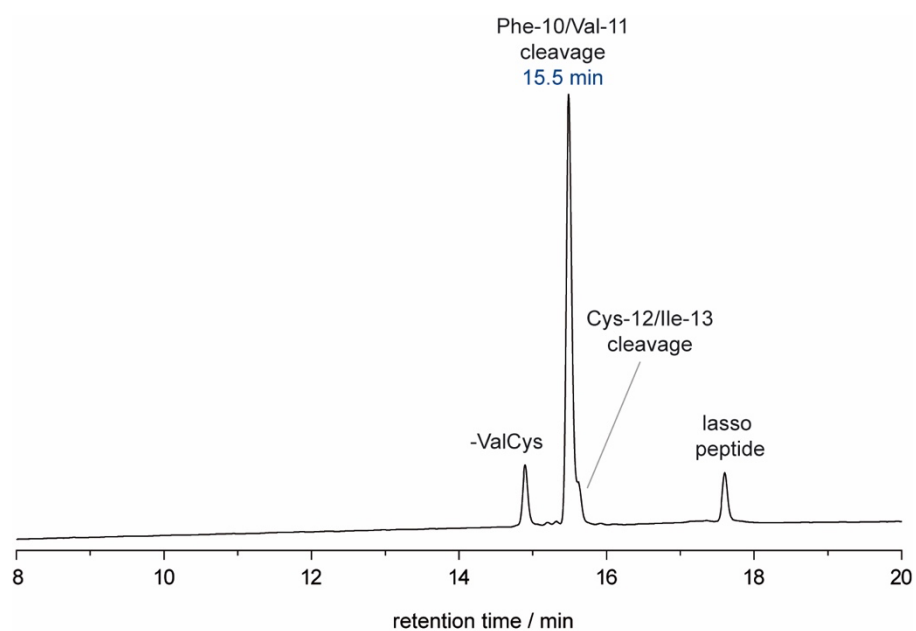
Figure 2. RP-HPLC chromatogram showing a thermolysin digestion of **L3** after 10 min at room temperature. The main product in the digest is the [2]rotaxane **3** after Phe-10/Val-11 cleavage.
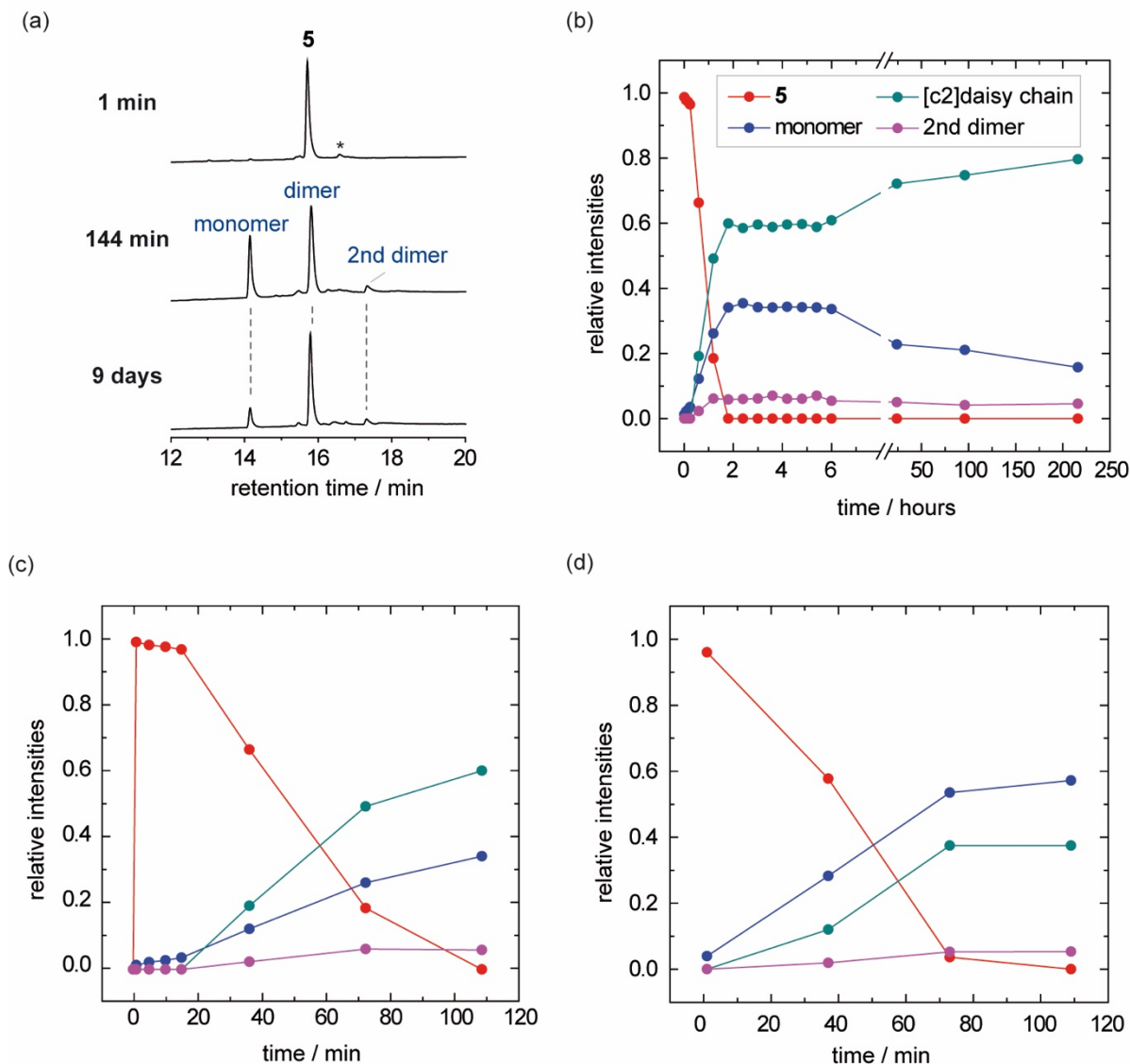
Figure 3. Kinetic profiles of formation of DCL (**5**-2H)$_n$ from lasso peptide **L5** (0.25 mM) in (NH$_4$)HCO$_3$ buffer (50 mM, pH 7.8). The relative intensities correspond to HPLC peak areas (UV absorbance at 215 nm) of different library members. (a) Representative HPLC chromatograms after 1 min, 144 min, and 9 days. (b) Kinetic profile over 9 days. (c) Zoom on the first 120 minutes (oxidation phase) in the self-assembly process. The time points < 20 min were taken after acidification of the reaction solution. (d) Self-assembly of **5** (0.25 mM) in the presence of 5 % (v/v) DMSO.

11
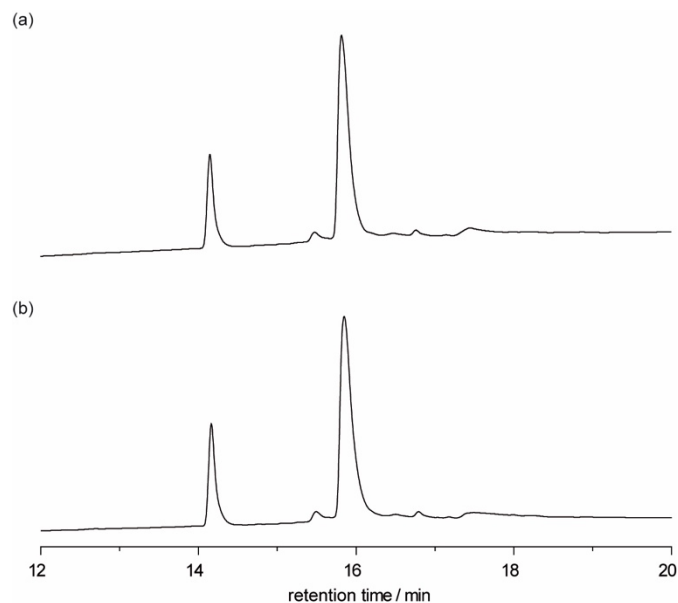
Figure 4. Fully oxidized DCL (**5**-2H)$_n$ (0.25 mM) in (NH$_4$)HCO$_3$ buffer (a) before and (b) after addition of 100 equivalents iodoacetamide. The mixture was left at room temperature for 2 h before it was analyzed by HPLC. No significant change was observed after addition of iodoacetamide. This indicates a full oxidation of the Cys residues.
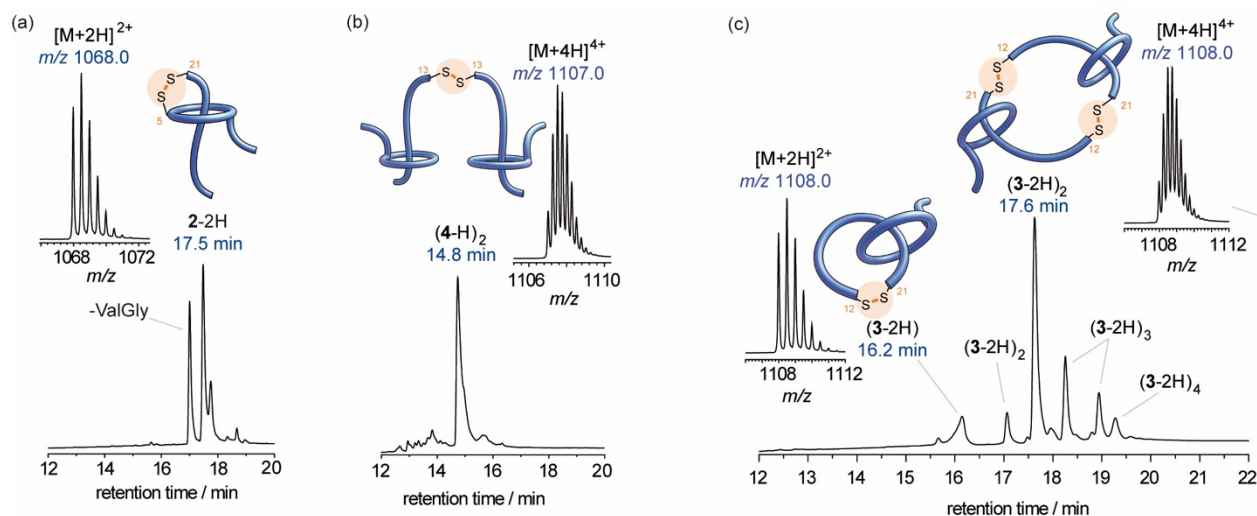
Figure 5. HPLC chromatograms and MS data of DCLs (0.25 mM) after self-assembly for 1 day in aqueous buffer: (a) (**2**-2H)$_n$ (Cys-5, Cys-21), (b) (**4**-H)$_n$ (Cys-13), and (c) (**3**-2H)$_n$ (Cys-12, Cys-21).
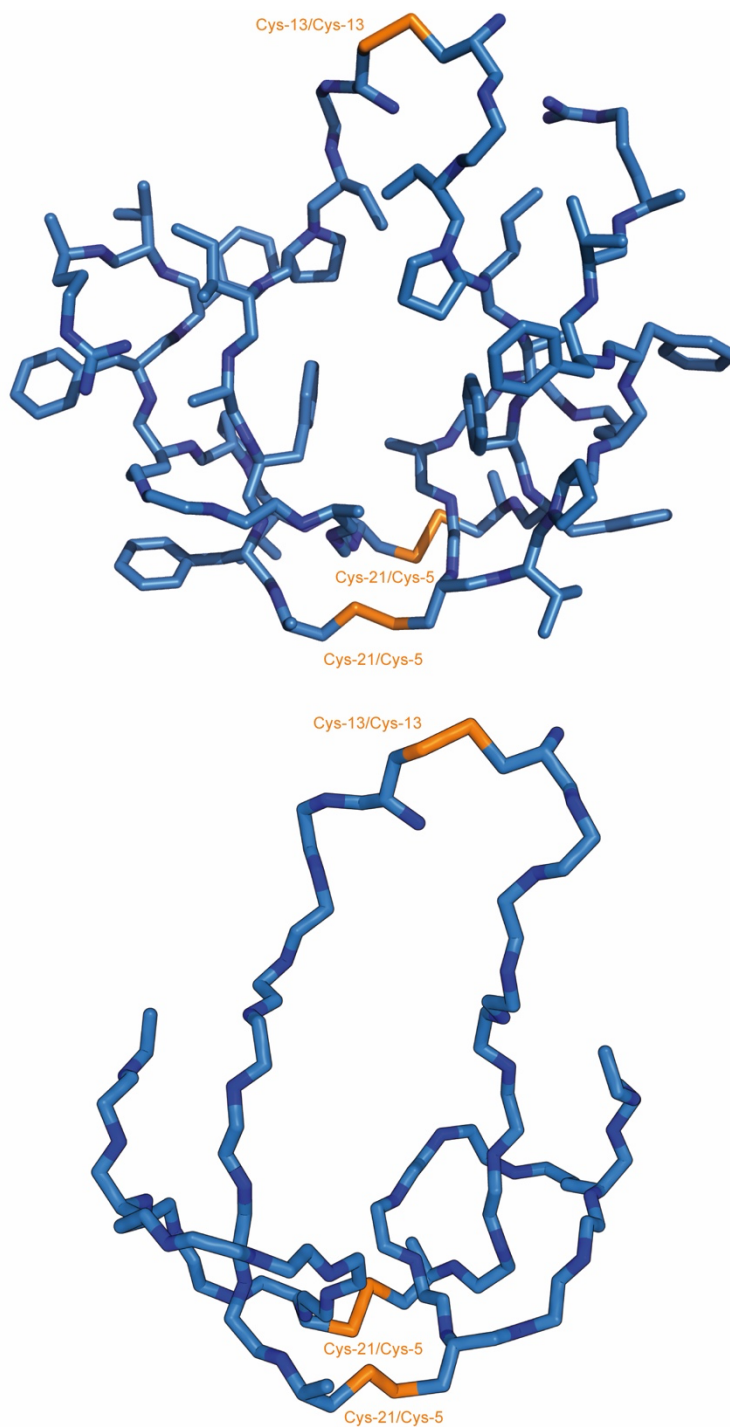
Figure 6. Energy-minimized structure of dimer (**7**-3H)$_2$ obtained by computational modeling at the GFN2-xTB[8] level of theory using conductor-like polarizable continuum solvent model[9] (water, ε = 80.4) in the ORCA[10] software package. The backbone structure is depicted as sticks with (top) and without (bottom) side chains. Hydrogen and oxygen atoms are not shown for the sake of clarity. Nitrogen atoms are shaded dark blue and the disulfide bonds (Cys-5/Cys-21, Cys-5/Cys-21 and Cys-13/Cys-13) are colored orange.
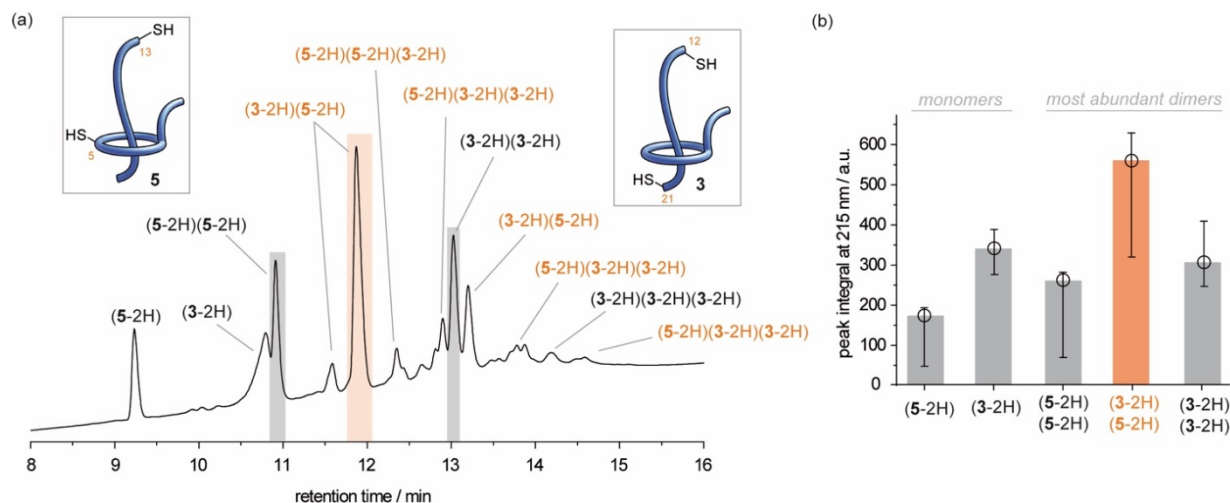
Figure 7. LC/MS analysis of mixed DCL with building blocks **3** (0.13 mM) and **5** (0.13 mM) at equimolar concentrations (0.25 mM overall building block concentration) in NH_4HCO_3 buffer (50 mM): (a) Chromatogram (UV absorbance at 215 nm) obtained by LC/MS after 2 days. (b) Peak area integrals correspond to the amount of monomers, homo- and heterodimers in the mixed DCL (**3**-2H)$_n$(**5**-2H)$_n$. The ratio of ~1:2:1 (homo-:hetero-:homodimer) reflects the statistically expected ratio. The experiments were carried out in triplicate and the error bars show the range of values.
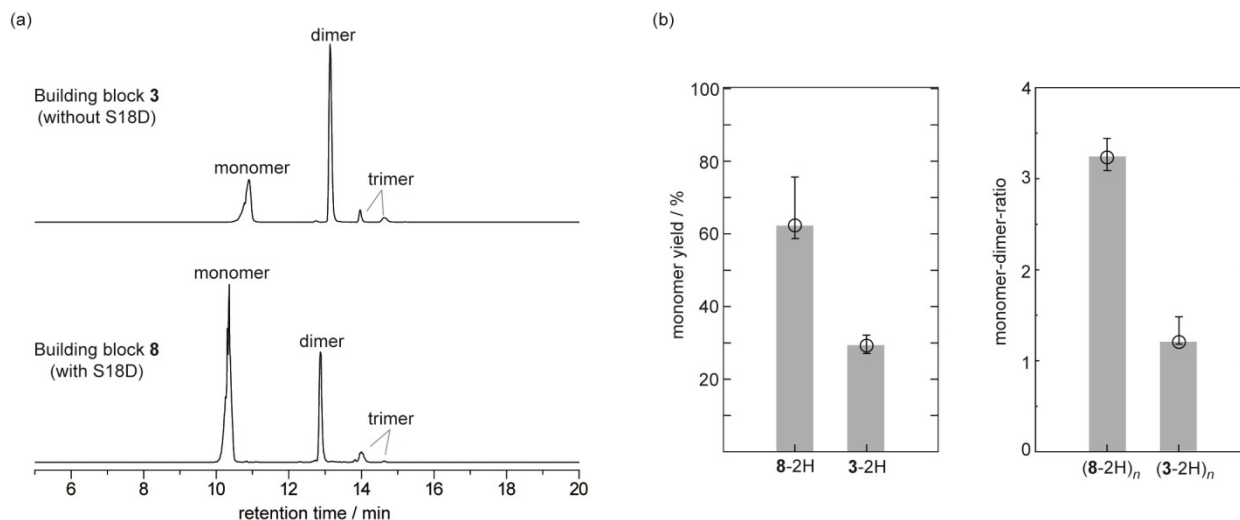
Figure 8. LC/MS analysis of DCLs (**3**-2H)$_n$ and (**8**-2H)$_n$ in $(NH_4)HCO_3$ buffer (50 mM) at a equimolar concentrations of 100 µM after 2 days. (a) Normalized extracted ion currents for (top) DCL (**3**-2H)$_n$ (*m/z* 1108) and (bottom) DCL (**8**-2H)$_n$ (*m/z* 1122). (b) Monomer yield and monomer-dimer-ration for both DCLs. Yields were determined from LC peak areas (UV absorbance at 215 nm). The experiments were carried out in triplicate and the error bars show the range of values.

14
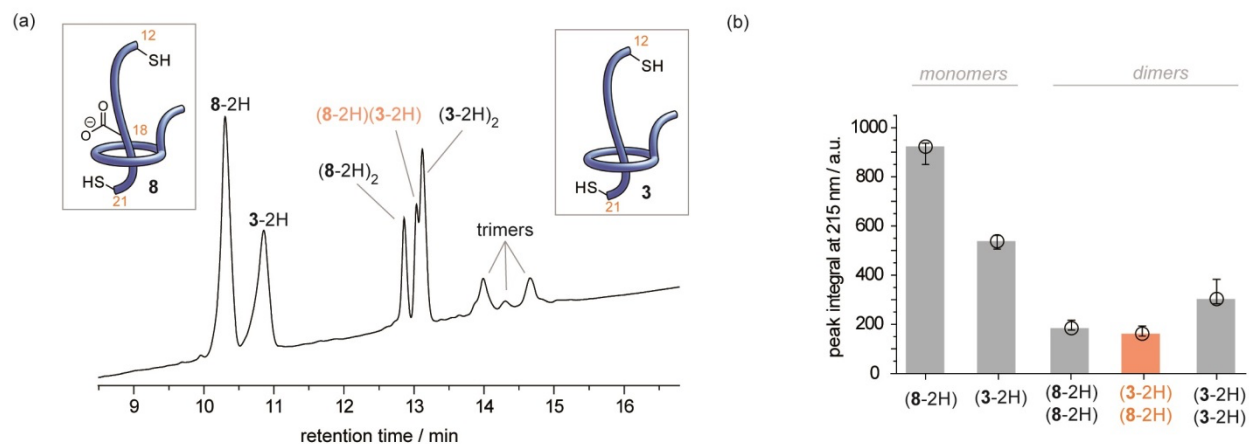
Figure 9. LC/MS analysis of mixed DCL containing building blocks **3** (50 μM) and **8** (50 μM) in $(NH_4)HCO_3$ buffer (50 mM) at an overall building block concentration of 100 μM after 2 days: (a) Chromatogram (UV absorbance at 215 nm) obtained by LC/MS and (b) LC peak area integrals correspond to the amount of monomers, homo- and heterodimers in the mixed DCL. The experiments were carried out in triplicate and the error bars show the range of values.
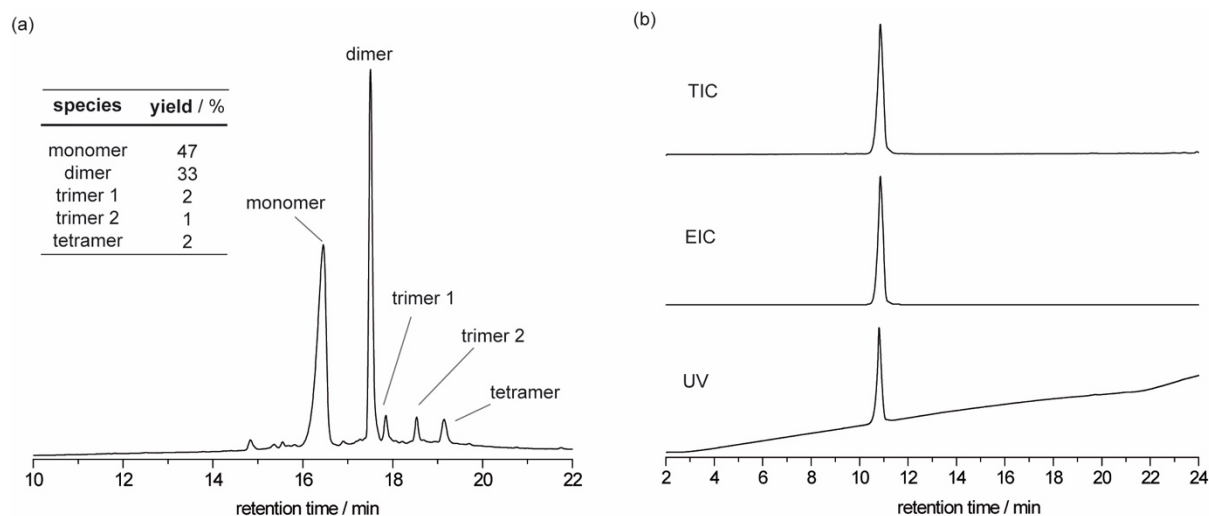
Figure 10. HPLC and LC/MS analysis of [2]catenane synthesis: (a) The HPLC chromatogram of DCL (**3**-2H)$_n$ (25 μM) after 2 days at 37 °C shows that the [2]catenane has formed in 47% yield based on HPLC peak areas. (b) LC/MS analysis of isolated [2]catenane **3**-2H (1.03 mM) in ultrapure water (pH ~6.0) after 9 days at room temperature and a short-term heating to 64 °C demonstrates the stability of the disulfide-bonded structure (TIC = total ion current, EIC = extracted ion current for *m/z* 1108.5 and UV absorbance at 215 nm).
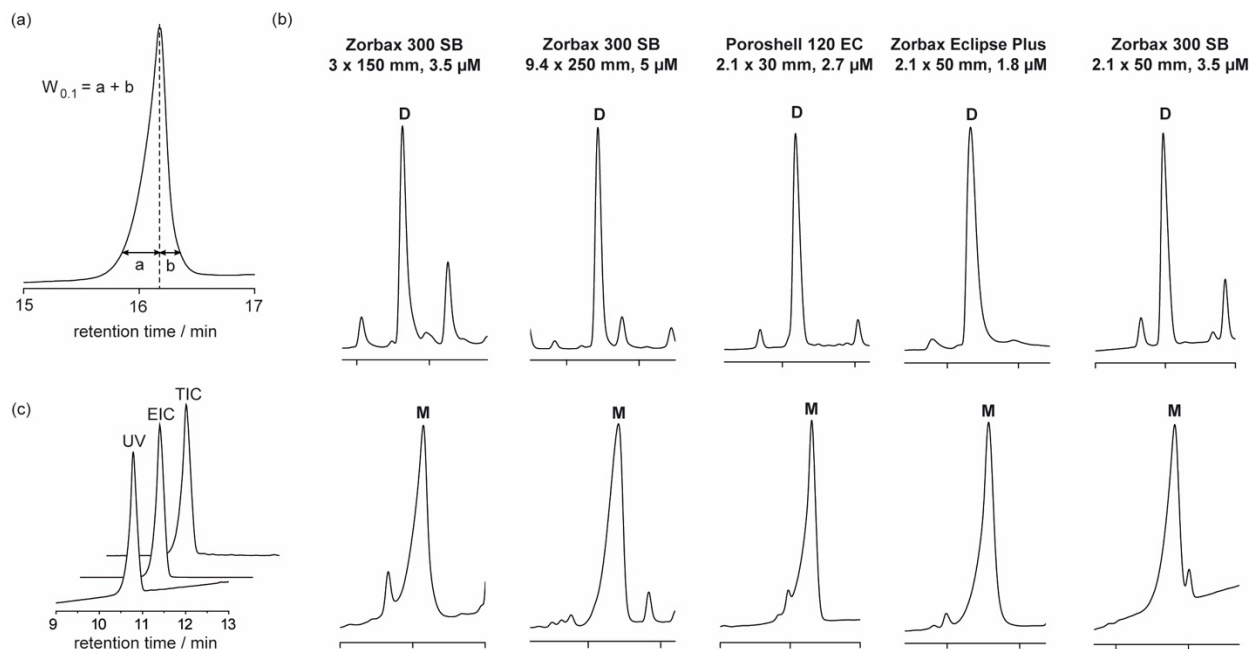
Figure 11. Peak fronting of [2]catenane **3**-2H on different RP-HPLC columns and peak shape of the dimer for direct comparison (see Table 6 below): (a) Injection of purified [2]catenane **3**-2H (5 µg) on a Zorbax 300 SB-C18 (3 x 150 mm, 3.5 µm) column. (b) Normalized monomer (M, [2]catenane **3**-2H) and dimer (D, [3]catenane (**3**-2H)$_2$) peaks from the same injection of DCL (**3**-2H)$_n$ on different RP-HPLC columns. (c) Injection of purified [2]catenane **3**-2H (2.5 µg) on a Zorbax 300 SB-C18 (2.1 x 50 mm, 3.5 µm) and corresponding MS experiment showing the absence of impurities (TIC = total ion current, EIC = extracted ion current for *m/z* 1108 and UV trace at 215 nm).

Table 6. Analysis of HPLC peak shape of [2]catenane **3**-2H and [3]catenane (**3**-2H)$_2$ on different reversed-phase columns

Column	Monomer peak				Dimer peak			
	Retention time / min	$W_{0.1}$ / min	a / min	b / min	Retention time / min	$W_{0.1}$ / min	a / min	b / min
Zorbax 300 SB-C18[a] 3 x 150 mm, 3.5 µm	16.2	0.49	0.33	0.16	17.6	0.21	0.07	0.14
Zorbax 300 SB-C18[b] 9.4 x 250 mm, 5 µm[a]	16.4	0.43	0.31	0.12	17.4	0.17	0.07	0.10
Poroshell 120 EC-C18 2.1 x 30 mm, 2.7 µm	13.3	0.40	0.28	0.12	15.18	0.19	0.07	0.12
Zorbax Eclipse Plus C18[c] 2.1 x 50 mm, 1.8 µm	20.9	0.45	0.28	0.17	22.3	0.28	0.09	0.19
Zorbax 300 SB-C18[d] 2.1 x 50 mm, 3.5 µm	10.8	0.36	0.21	0.15	13.0	0.19	0.08	0.11

[a] DCL (10 µg) using a binary water/acetonitrile gradient with TFA (0.1%) and a flow rate of 0.75 mL/min

[b] DCL (50 µg) using a binary water/acetonitrile gradient with TFA (0.1%) and a flow rate of 4.0 mL/min

[c] DCL (10 µg) using a binary water/acetonitrile gradient with TFA (0.1%) and a flow rate of 0.3 mL/min

[d] DCL (5 µg) using a binary water/acetonitrile gradient with formic acid (0.1%) and a flow rate of 0.5 mL/min
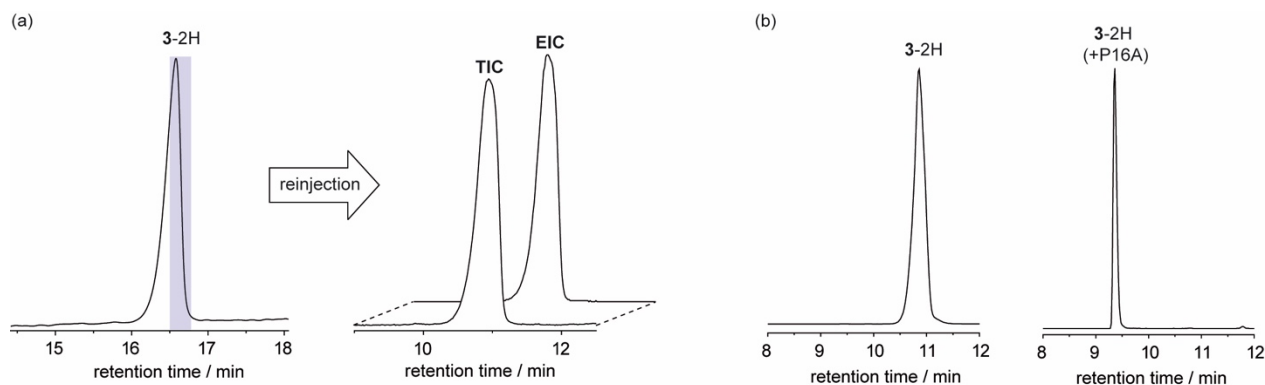
Figure 12. (a) Reinjection of the HPLC-isolated trailing edge of the [2]catenane **3**-2H peak (left) generates the initially observed fronting as shown by subsequent LC/MS analysis (right) after 30 min at room temperature (TIC = total ion current, EIC = extracted ion current for *m/z* 1108). (b) Comparison between normalized LC/MS traces of [2]catenane **3**-2H (EIC for *m/z* 1108) and a structurally similar [2]catenane **3**-2H **P16A** with a substitution of Pro-16 by Ala (Cys-12, Cys-21, Ala-16; EIC for *m/z* 1095).
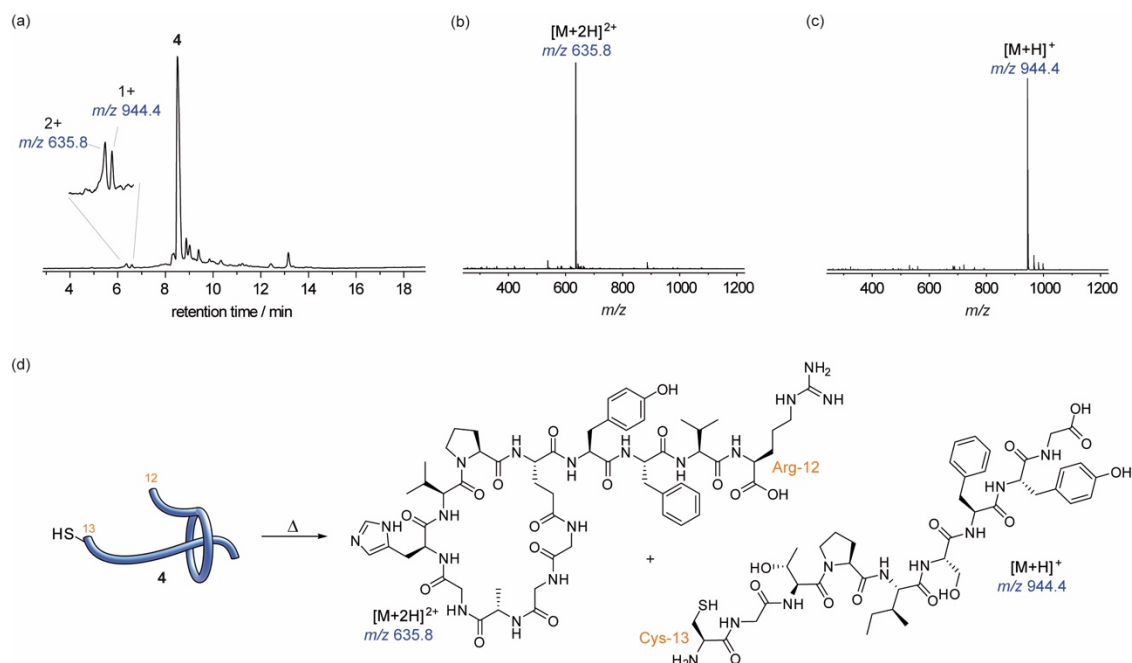
Figure 13. Heating experiment (95 °C, 2 h) with [2]rotaxane **4** in water containing DTT (5 mM) and formic acid (0.5 % v/v): (a) Chromatogram (TIC) after 2 h. Mass spectra of (b) peak of the isopeptide-bonded wheel (GGAGHVPEYFVR) and (c) the free axle molecule (CGTPISFYG) after disassembly of [2]rotaxane **4**. (d) Schematic representation of heat-driven unthreading of [2]rotaxane **4**. The stability of **4** and the slow rate of unthreading demonstrate that [2]rotaxane **4** is stable and that the peptide wheel is tightly held in position by the bulky stopper residues Phe-19 and Tyr-20, even at elevated temperature.
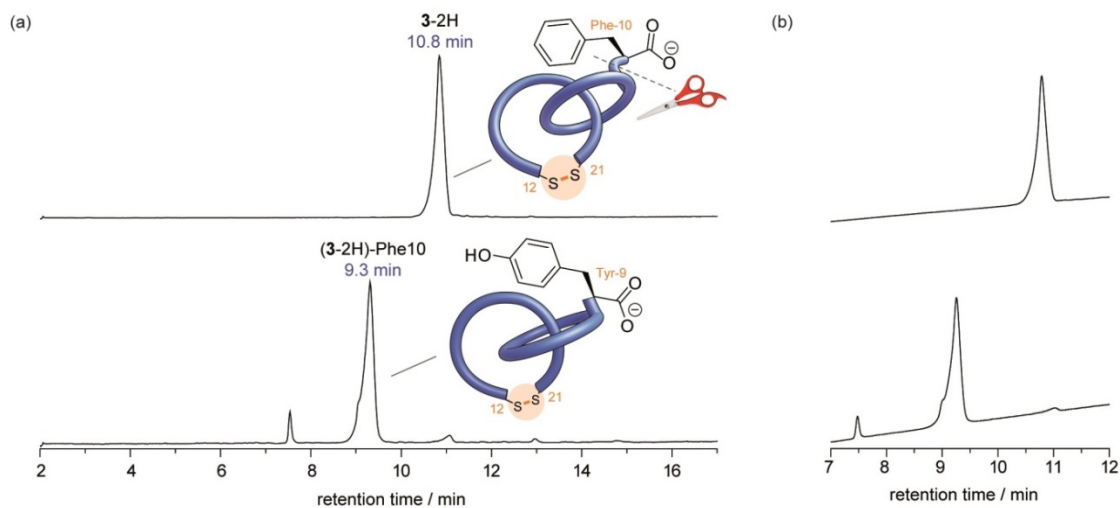
Figure 14. LC/MS analysis of carboxypeptidase assay of [2]catenane **3**-2H: (a) Total ion currents of [2]catenane **3**-2H before (top) and after carboxypeptidase digestion for 31 h (bottom) and (b) corresponding UV traces (215 nm absorbance).

18

5. Structure Analysis and Assignment of Disulfide Connectivities by MS²/CID

All MccJ25 variants in this study display a similar fragmentation pattern, which involves energetically preferred cleavage between Val-6/Pro-7, Thr-15/Pro-16, and Ile-17/Ser-18 (Figure 15). This effect related to Pro (so-called proline effect) and other residues during gas phase fragmentation via CID of peptides was reported previously.[11] Thus, a protocol was applied in which the ion of interest was mass-selected and subsequently fragmented at different fragmentation voltages. Low energy fragmentation pathways are initially frequented yielding large peptide fragments instead of a mixture of small fragments. Analysis of these larger fragment ions allows for a re-construction of the initial structure and topology of a MIP. An additional fragmentation pattern is observed for the isopeptide-bonded macrocycle (Figure 15e): low-energy fragmentations generate b- and y-ions for residues not involved in the macrocycle (e.g. Phe-10). The b_8-ion undergoes a subsequent CO loss yielding the corresponding immonium species (a-ions).
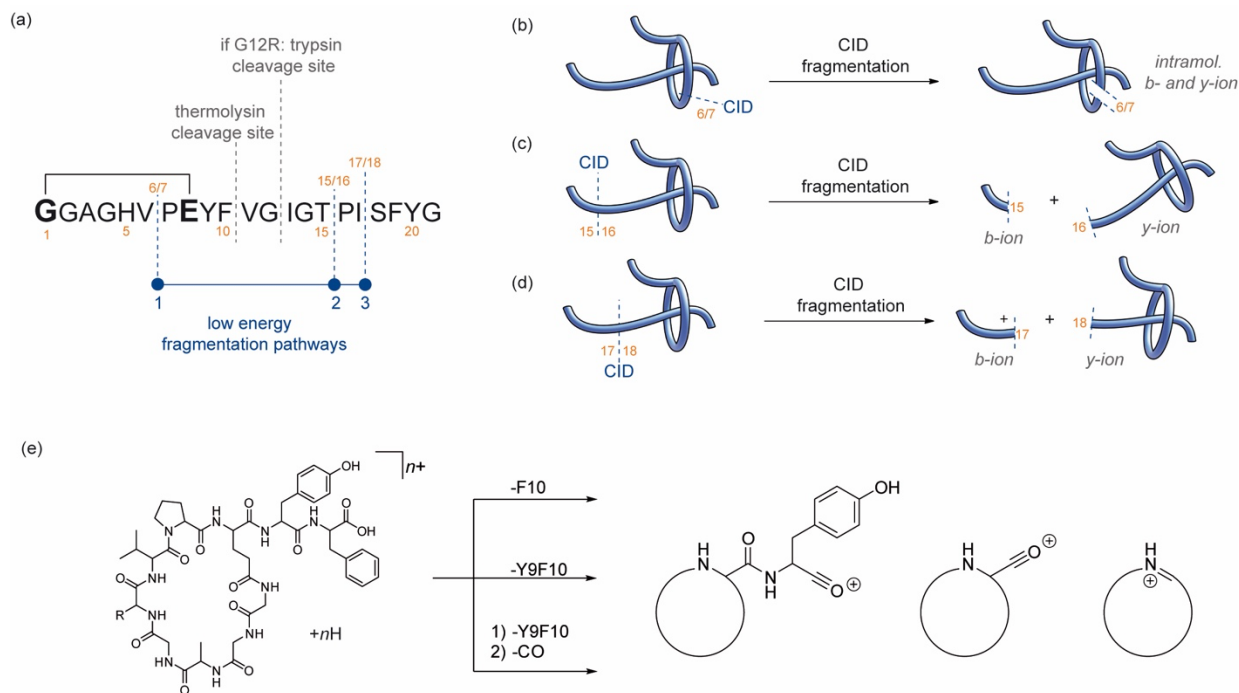
Figure 15. Schematic representation of low-energy fragmentation reactions induced by CID, which help to identify different MIP topologies and structures: (a) Wild-type sequence of MccJ25 with enzymatic cleavage sites. (b-d) Different gas phase fragmentations of MIPs. The dotted blue lines indiate low energy fragmentation pathways during CID which yield b- and y-ions of the peptide constructs. (e) Gas phase fragmentation of isopeptide-bonded macrocycle.

Structure of Building Block 3
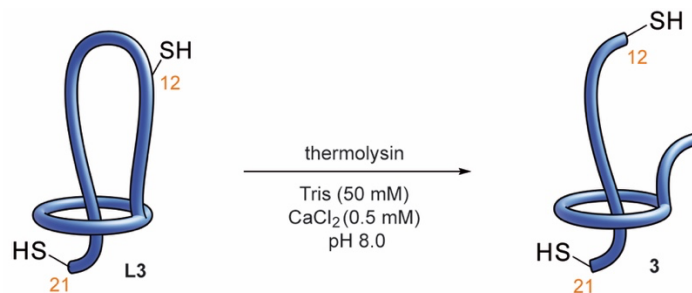
Figure 16. Schematic representation of the thermolysin digestion of **L3**.
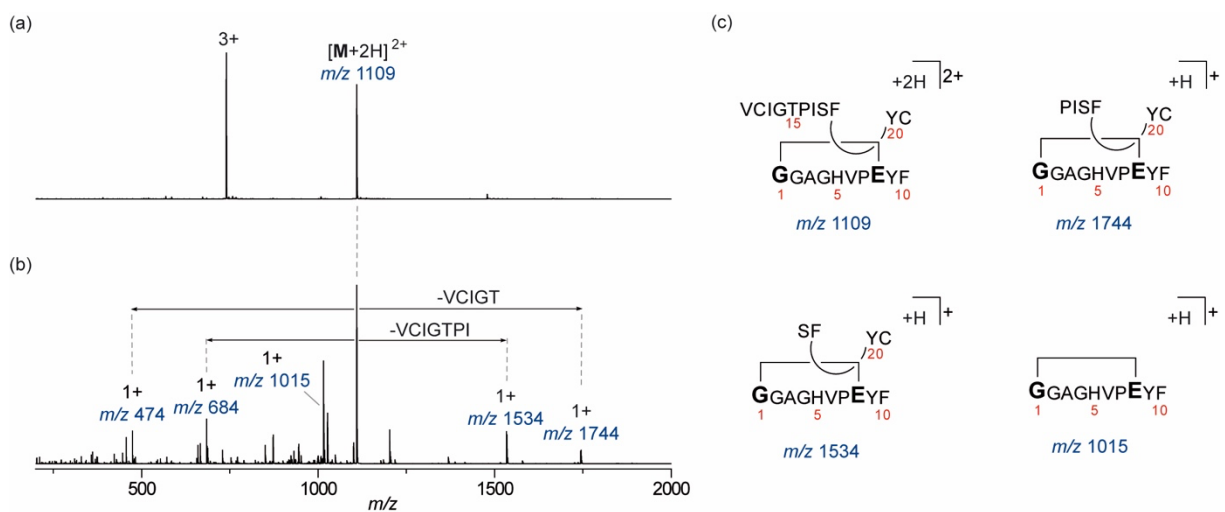
Figure 17. Tandem mass spectrometry (MS^2) with collision-induced dissociation (CID) of [2]rotaxane **3**: ESI mass spectrum (a) before and (b) after mass selection of *m/z* 1109 and fragmentation using a collision voltage of 40 V. (c) Structures of observed fragment ions. Signals at *m/z* 1534 and 1744 indicate a loss of the fragments VCIGTPI and VCIGT, respectively. Thus, a thermolysin digestion cleaves variant **L3** initally between Phe-10 and Val-11.

Structure and Disulfide Connectivities of [c2]Daisy Chain (5-2H)₂
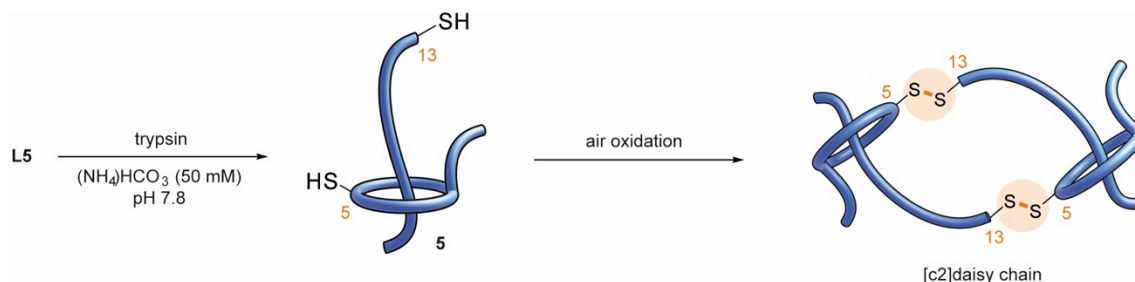
Figure 18. Schematic representation of the self-assembly synthesis of [c2]daisy chain (**5**-2H)₂ from lasso peptide **L5**.
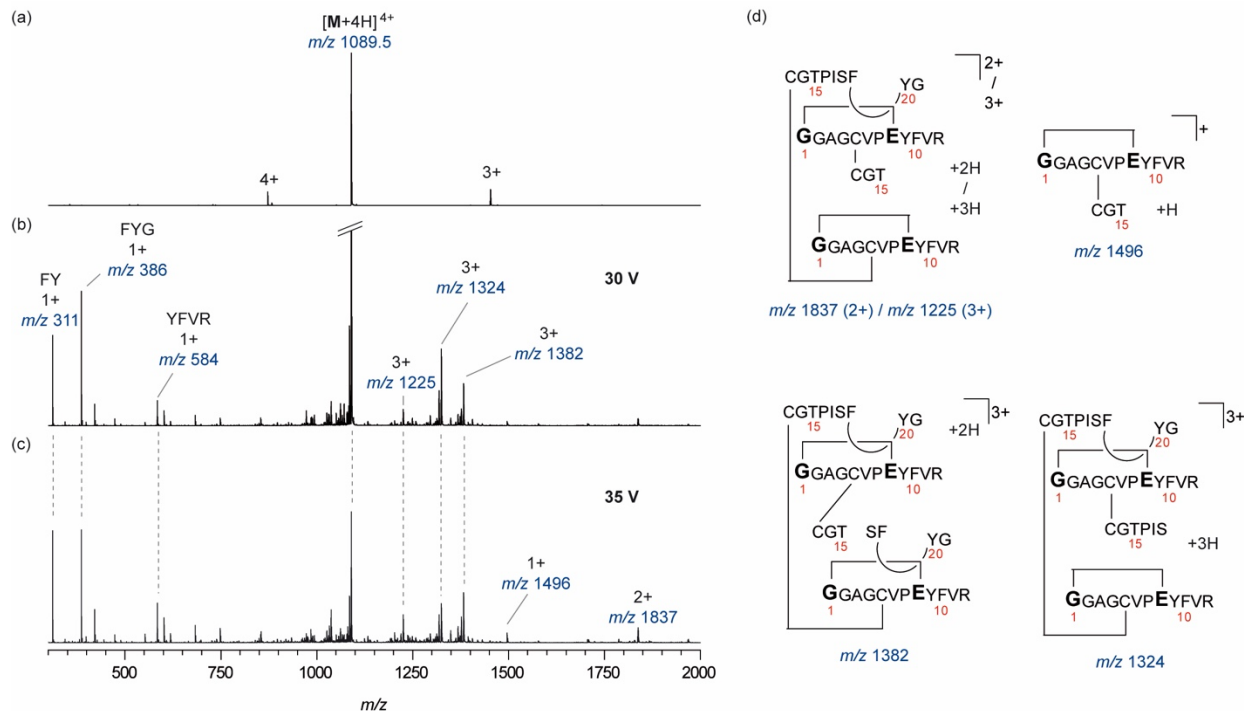
Figure 19. CID experiments with [c2]daisy chain (**5**-2H)₂: (a) ESI mass spectrum of (**5**-2H)₂. (b) Mass selection of *m/z* 1089.5 and CID experiments with (b) 30 V and (c) 35 V collision voltage. (d) Suggested chemical structures for observed *m/z* values, which indicate a "head-to-tail" instead of a "head-to-head" structure for the dimer. For example, the fragment corresponding to *m/z* 1496 can only be generated by a disulfide bond between Cys-5 (ring) and Cys-13 (thread).

Structure and Disulfide Connectivities of [1]Rotaxane 2-2H
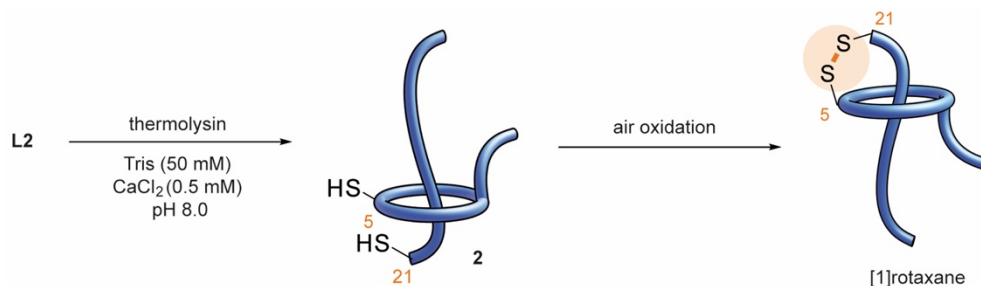
Figure 20. Schematic representation of the self-assembly synthesis of [1]rotaxane **2**-2H from lasso peptide **L2**.
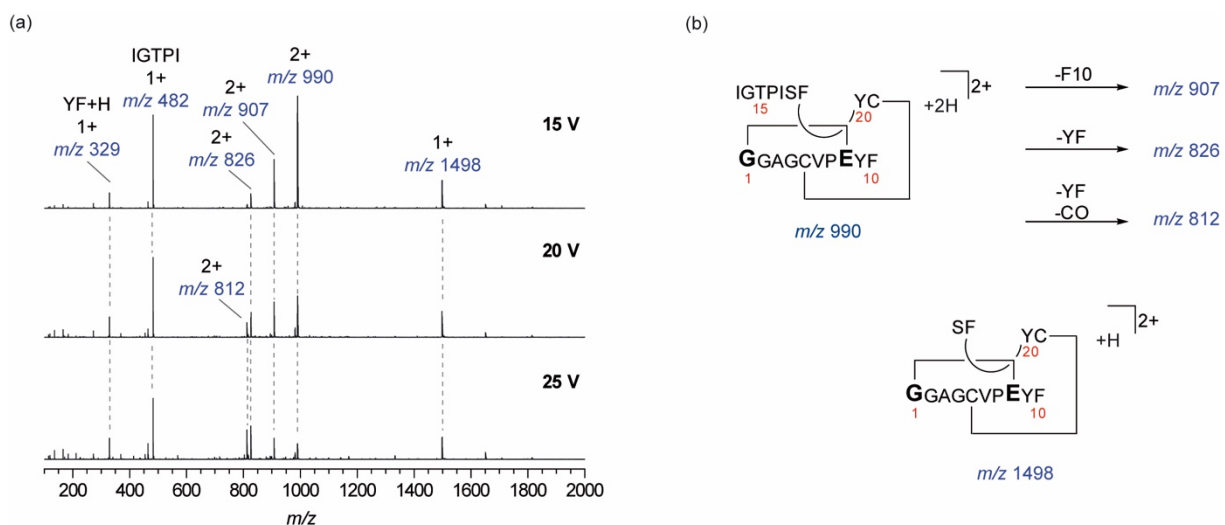
Figure 21. MS^2/CID experiments with [1]rotaxane **2**-2H **ΔN2** (monomer) after loss of N-terminal ValGly residue by a second thermolysin cleavage: (a) mass spectra at different collision voltages after mass-selection of [M+2H]$^{2+}$ (*m/z* 990). (b) Chemical structures for the observed fragments in the CID experiment.

Structure and Disulfide Connectivities of [3]Rotaxane (4-2H)₂
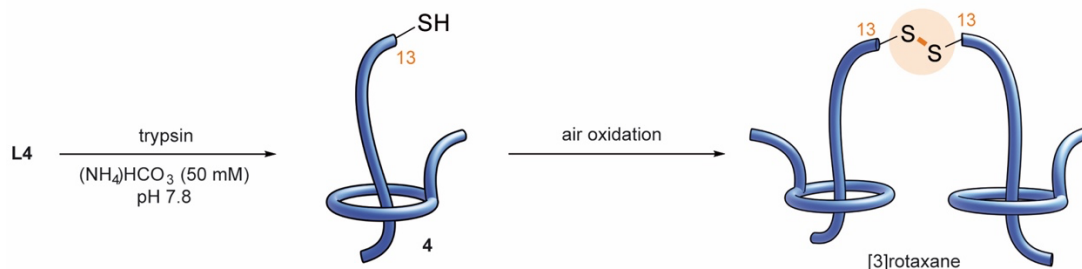
Figure 22. Schematic representation of the self-assembly synthesis of [3]rotaxane (**4**-2H)₂ from lasso peptide **L4**.
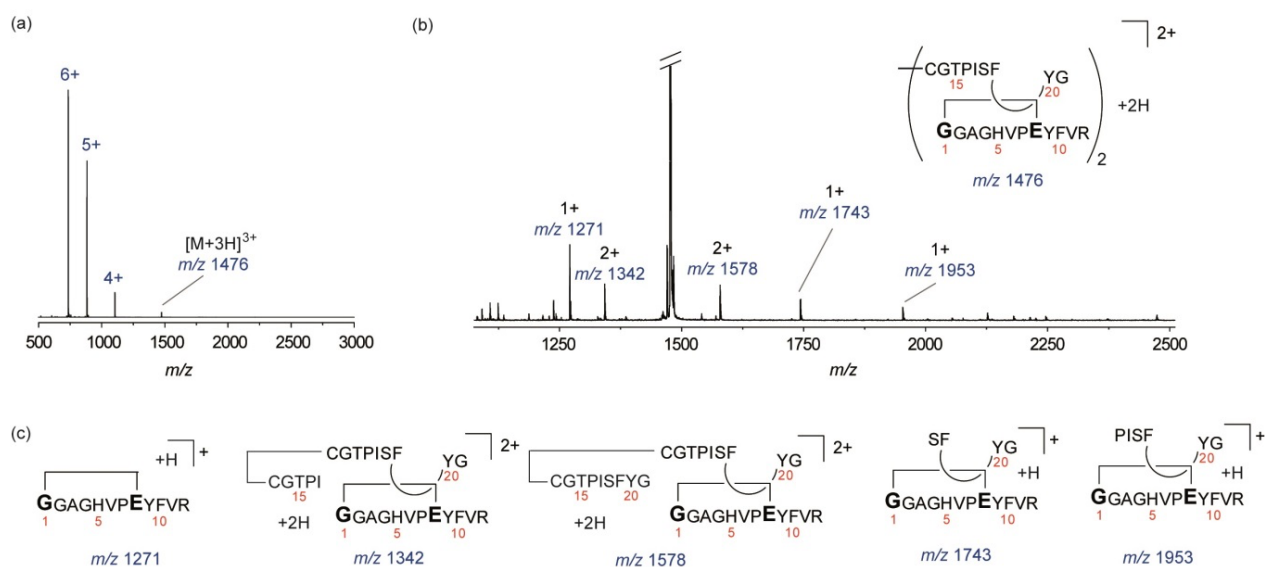
Figure 23. CID experiments with [3]rotaxane (**4**-H)₂: (a) ESI mass spectrum of (**4**-H)₂. (b) Mass selection of *m/z* 1476 and CID experiment (45 V collision voltage). The CID experiment shows the initially formed fragments, *m/z* 1342, 1578, 1743, and 1953, correspond to [2]rotaxanes after scission of the symmetric axle molecule. (c) Chemical structures for the observed key fragments, which confirm the [3]rotaxane structure of the dimer.

Structure and Disulfide Connectivities of [*n*+1]Catenanes (3-2H)$_n$
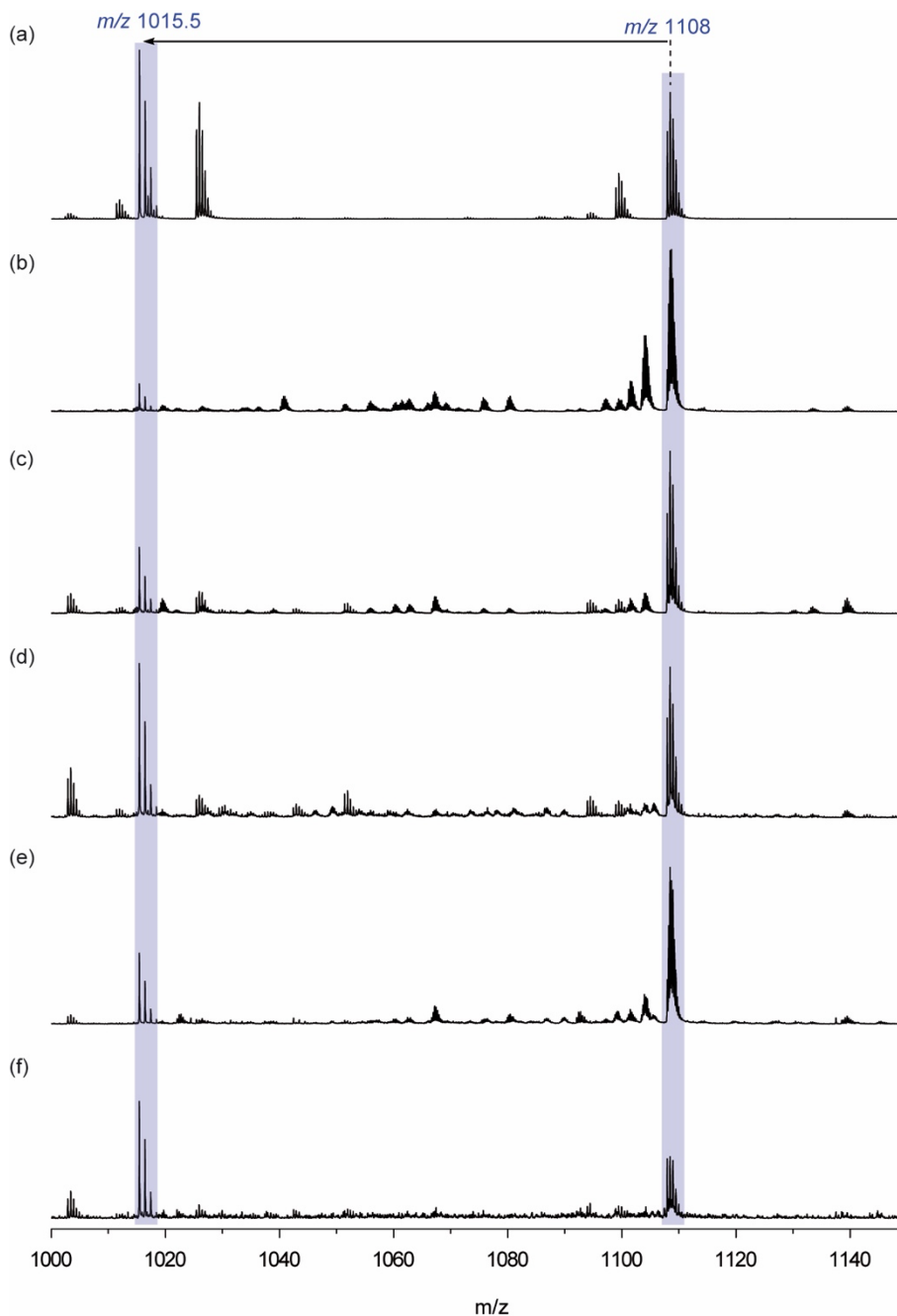
Figure 24. CID experiments (30 V fragmentation voltage) with library members of DCL (**3**-2H)$_n$ (see Figure 5): (a) monomer (16.2 min HPLC retention time), (b) head-to-head dimer (17.1 min), (c) head-to-tail dimer (17.6 min), (d) trimer (18.3 min), (e) trimer (18.9 min) and (f) tetramer (19.3 min). Loss of the fragmented, isopeptide-bonded wheel (*m/z* 1015.5) in the gas phase indicates a non-covalently bound structure and, thus, a [*n*+1]catenane topology for all library members.
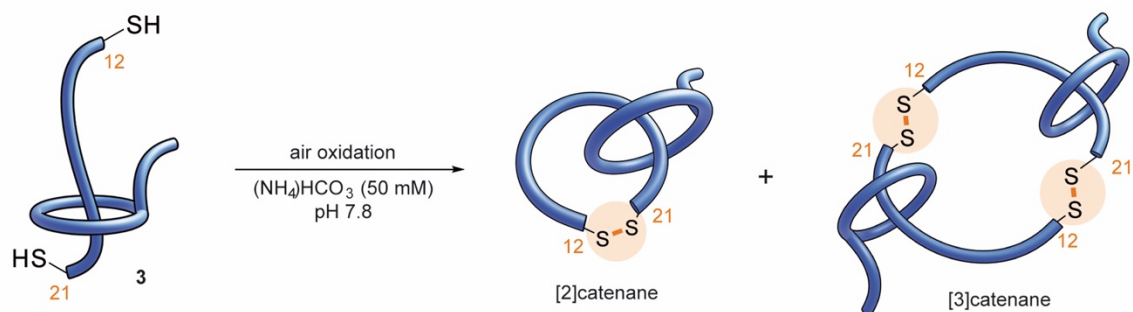
Figure 25. Schematic representation of the self-assembly synthesis of [2-3]catenanes (**3**-2H)$_{2-3}$ from building block **3**.
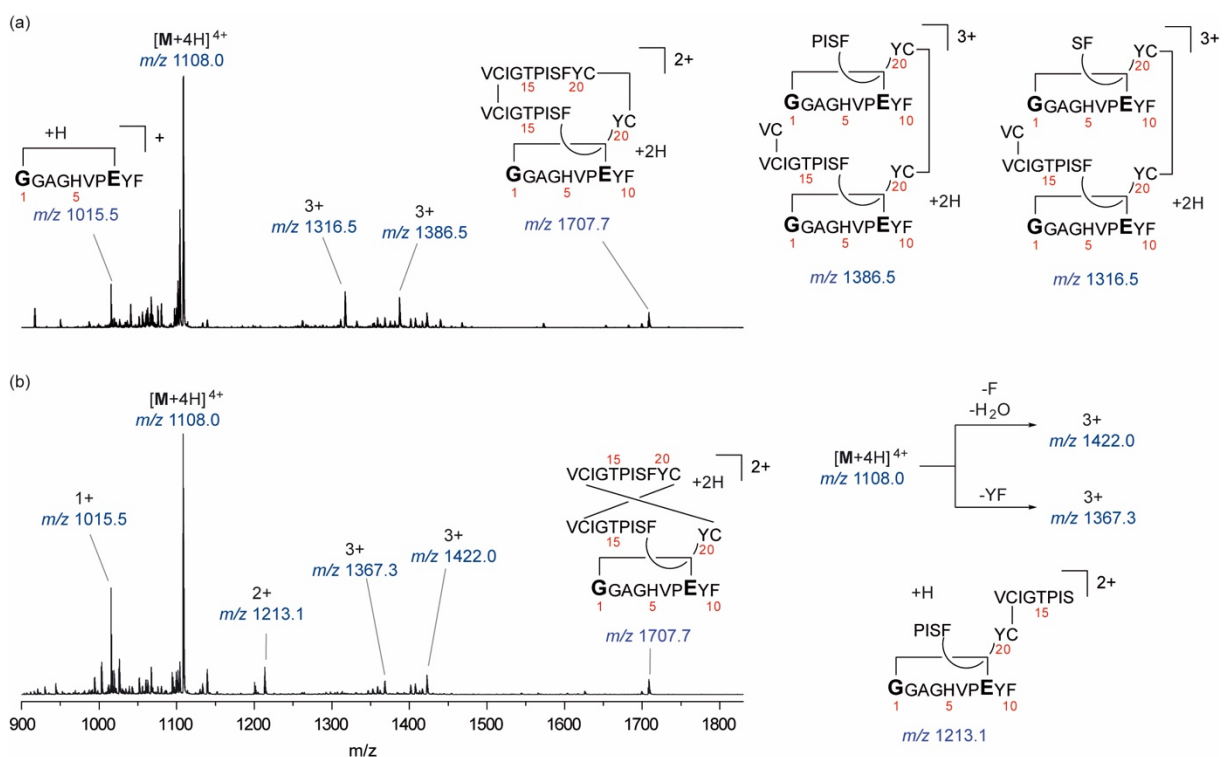
Figure 26. CID experiments (30 V fragmentation voltage) with [3]catenanes, (a) the head-to-head (17.1 min HPLC retention time) and (b) head-to-tail dimer (17.6 min) of DCL (**3**-2H)$_n$, and corresponding chemical structures for the observed key fragments which indicate the disulfide connectivity.
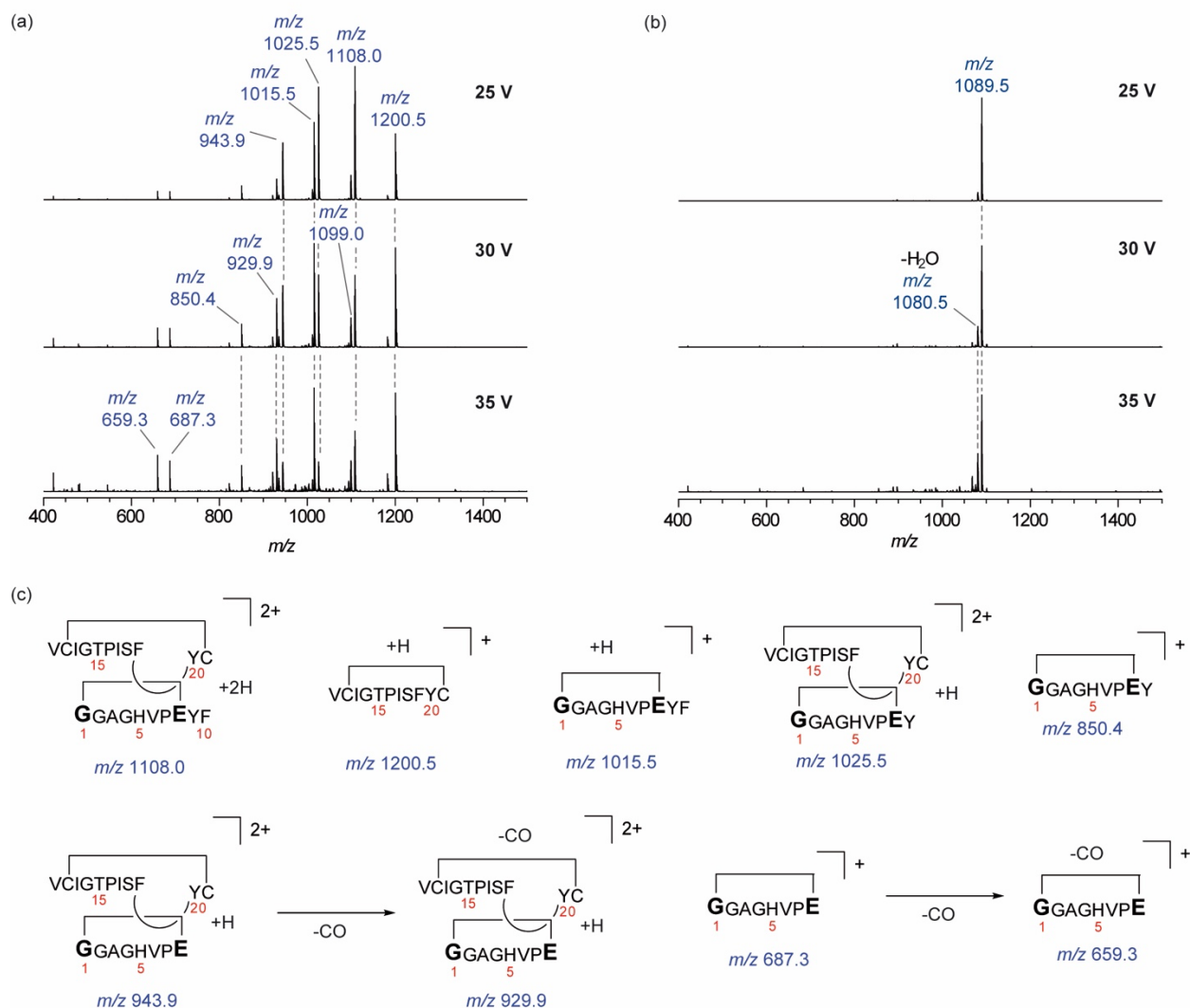
Figure 27. CID experiments with [2]catenane **3**-2H: (a) Fragmentation of mass-selected [M+2H]$^{2+}$ ion (*m/z* 1108) at different CID voltages. (b) For comparison, the same experiment with monomer **5**-2H at identical CID voltages shows minor fragmentation indicating that both rings of [2]catenane **3**-2H are not covalently bound and that the structure dissassembles more easily in the gas phase after first amide bond cleavage. (c) Chemical structures for the observed fragments in the CID experiment for [2]catenane **3**-2H.
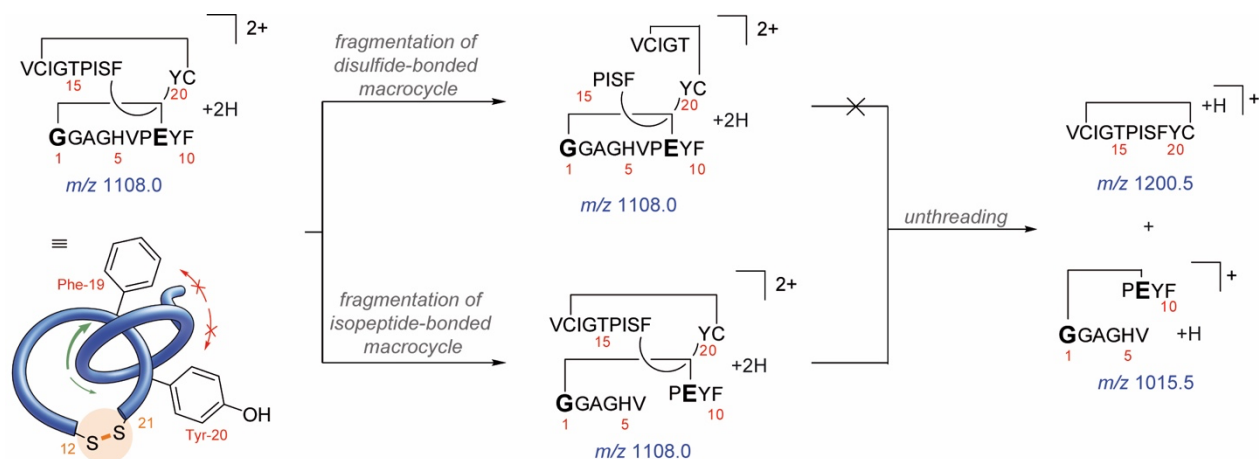
Figure 28. Suggested mechanism for gas phase fragmentation of [2]catenane **3**-2H. The disassembly of the disulfide-bonded ring and the fragment of the initially cleaved isopeptide-bonded ring indicates that at least partial through-the-annulus pirouetting is possible. A simultaneous double cleavage (cleavage of both rings) is energetically disfavored as shown in the example of monomer **5**-2H (Figure 27b).
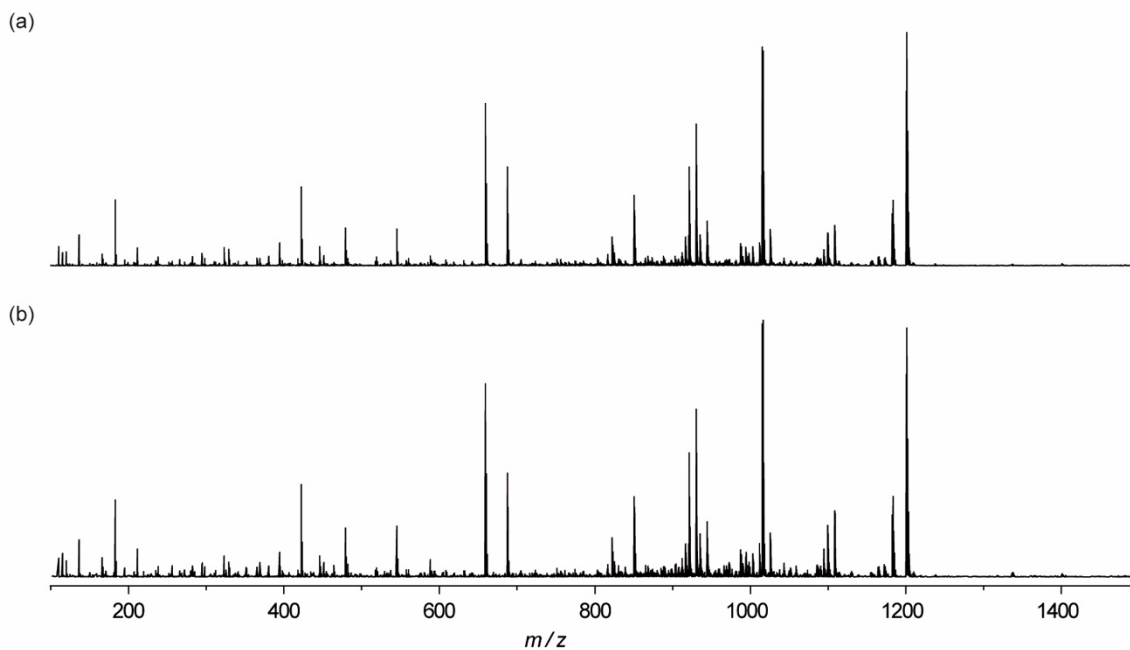
Figure 29. CID experiments with mass-selected [2]catenane **3**-2H (*m/z* 1108) at 40 V fragmentation voltage: (a) leading (10.5–10.7 min) and (b) trailing edge (10.9–11.1 min) of the [2]catenane peak during LC purification (see Figure 11). Assuming the presence of interconverting species, the very similar mass spectra indicate a similar fragmentation mechanism and, thus, structurally related species such as co-conformers. The presence of structural isomers would lead to significantly different fragmentations in the CID experiment.

Structure and Disulfide Connectivities of [2]Rotaxane 6-2H and [3]Rotaxane (6-3H)₂
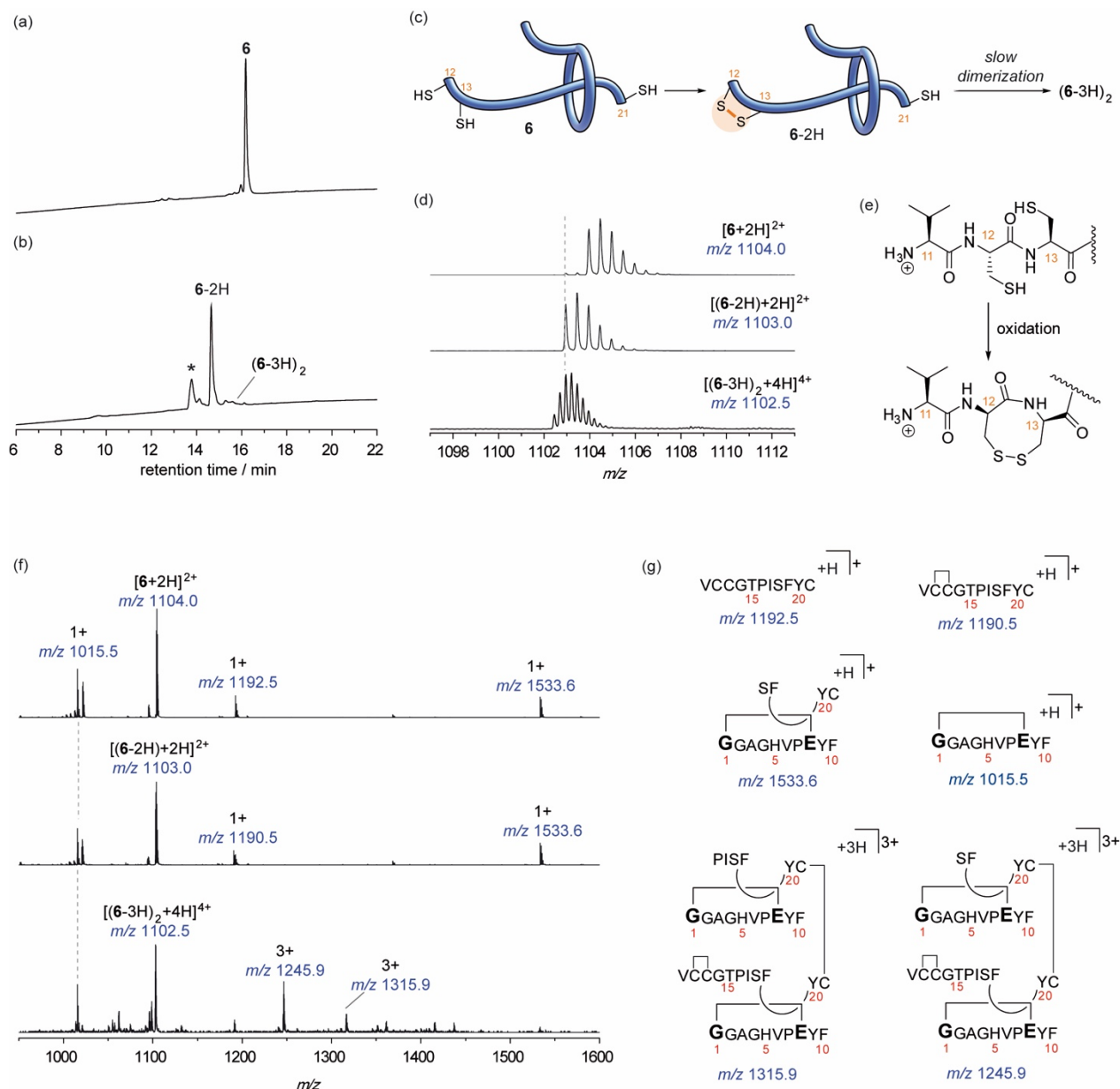
Figure 30. (a) HPLC chromatogram of DTT-reduced building block **6** before oxidation. (b) HPLC chromatogram after 1 day in aqueous buffer (pH 8.0). (c) Schematic representation of intramolecular disulfide formation (Cys-12/Cys-13) followed by a slow intermolecular dimerization (Cys-21/Cys-21). (d) ESI mass spectra of MIPs **6**, **6**-2H and (**6**-3H)₂. Only traces of the dimer were detected by LC/MS after 1 day. (e) First oxidation of building block **6** and formation of an intramolecular disulfide bond between Cys-12 and Cys-13. (f) CID experiments (30 V fragmentation voltage) with building block **6** (top), [2]rotaxane **6**-2H (middle) and [3]rotaxane (**6**-3H)₂ (bottom). (g) Chemical structures for the observed key fragments in the CID experiments.

Structure and Disulfide Connectivities of Double- and Tetra-Lasso Macrocycle (7-3H)₂ and (7-3H)₄
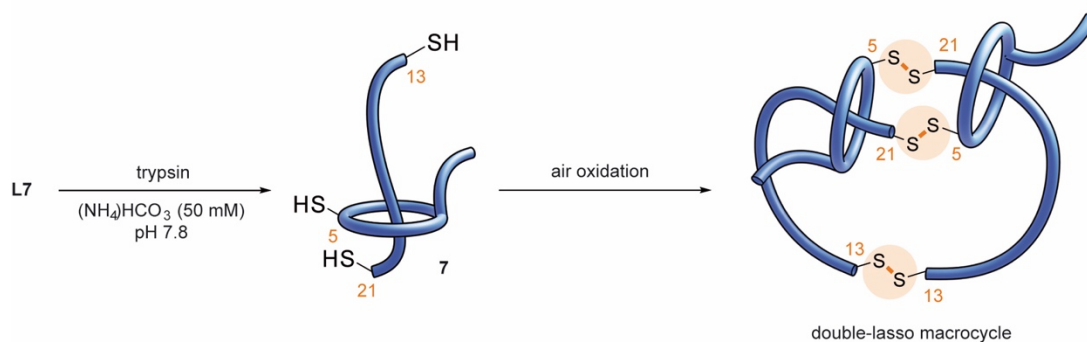
Figure 31. Schematic representation of the self-assembly synthesis of double-lasso macrocycle (**7**-3H)₂ from lasso peptide **L7**.
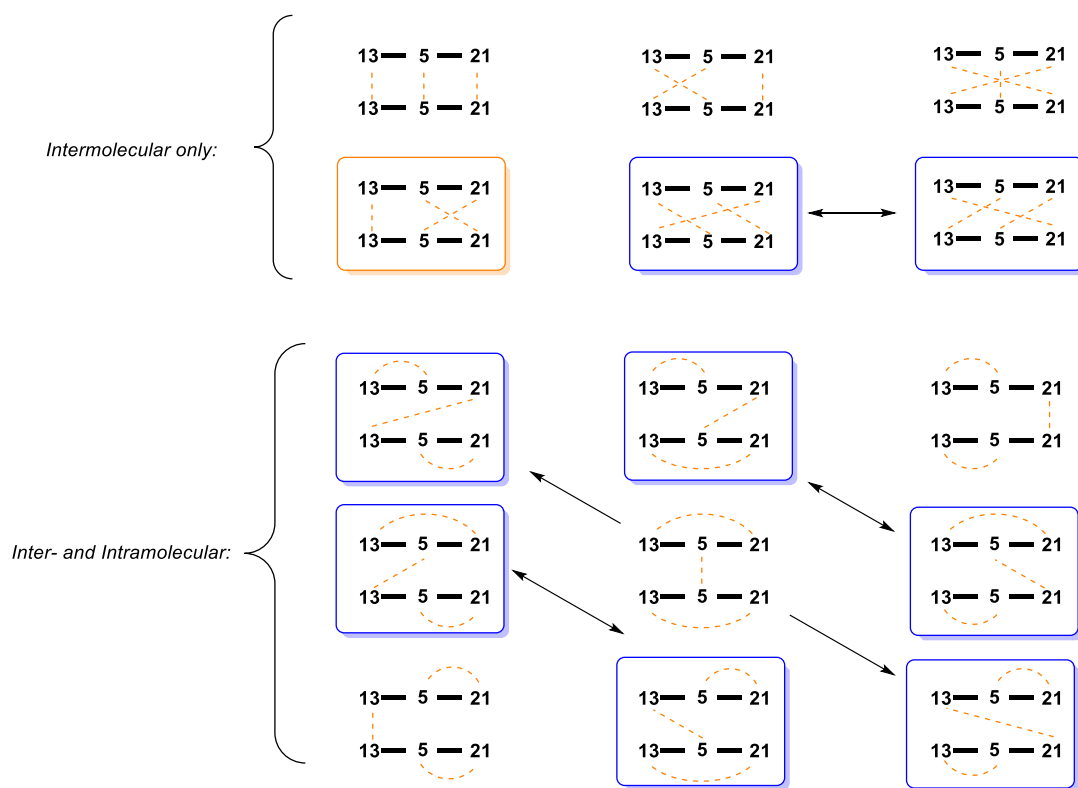
Figure 32. Possible disulfide connectivities in a homodimer with 6 cysteine residues: 4 (blue box and arrows) of the initial 15 species can be transferred into 4 others by rotation. The orange box marks the observed disulfide connectivity. Different molecular topologies arising by the spatial arrangement of crossing disulfide bonds are not taken into account.
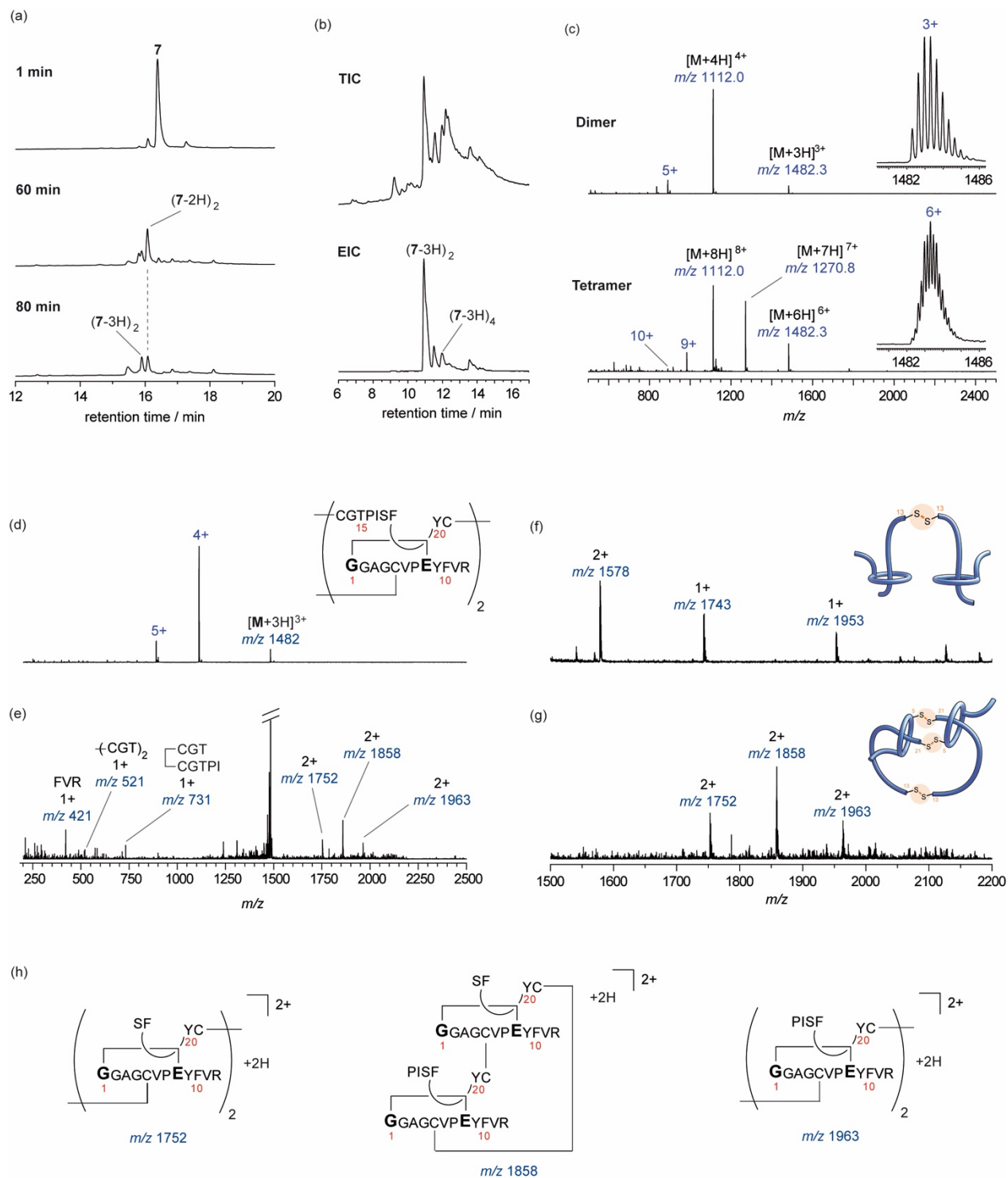
Figure 33. (a) HPLC chromatograms of the self-assembly of **7** show the instability of the DCL over time. (b) A LC/MS measurement after 90 min demonstrates that the dimer (**7**-3H)$_2$ is the most abundant consitutent in the DCL (TIC = total ion current, EIC = extracted ion current for *m/z* 1112). (c) ESI-MS spectra for the dimer and the tetramer with isotopic patterns of the 3+ and 6+ species, respectively. CID experiments with dimer (**7**-3H)$_2$: (d) ESI mass spectrum of (**7**-3H)$_2$.

30

(e) Mass selection of *m/z* 1482 and its fragmentation using a collision voltage of 55 V. The fragments at *m/z* 521 and 731 directly indicate a Cys-13/Cys-13 connectivity. (f) For comparison, CID experiment (45 V collision voltage) with linear [3]rotaxane $[(4\text{-H})_2+3\text{H}]^{3+}$ shows that initial fragmentation splits the [3]rotaxane into [2]rotaxane fragments (*m/z* 1578, 1743, and 1953) as shown in Figure 23 in detail. (g) CID experiment (55 V collision voltage) with $(7\text{-}3\text{H})_2$ proceeds similarly compared to [3]rotaxane $[(4\text{-H})_2+3\text{H}]^{3+}$ with axle scisson between Thr-15/Pro-16 and Ile-17/Ser-18 residues. (h) Instead of [2]rotaxane fragments, however, larger fragments are generated due to the presence of three disulfide bonds connecting the two building blocks.
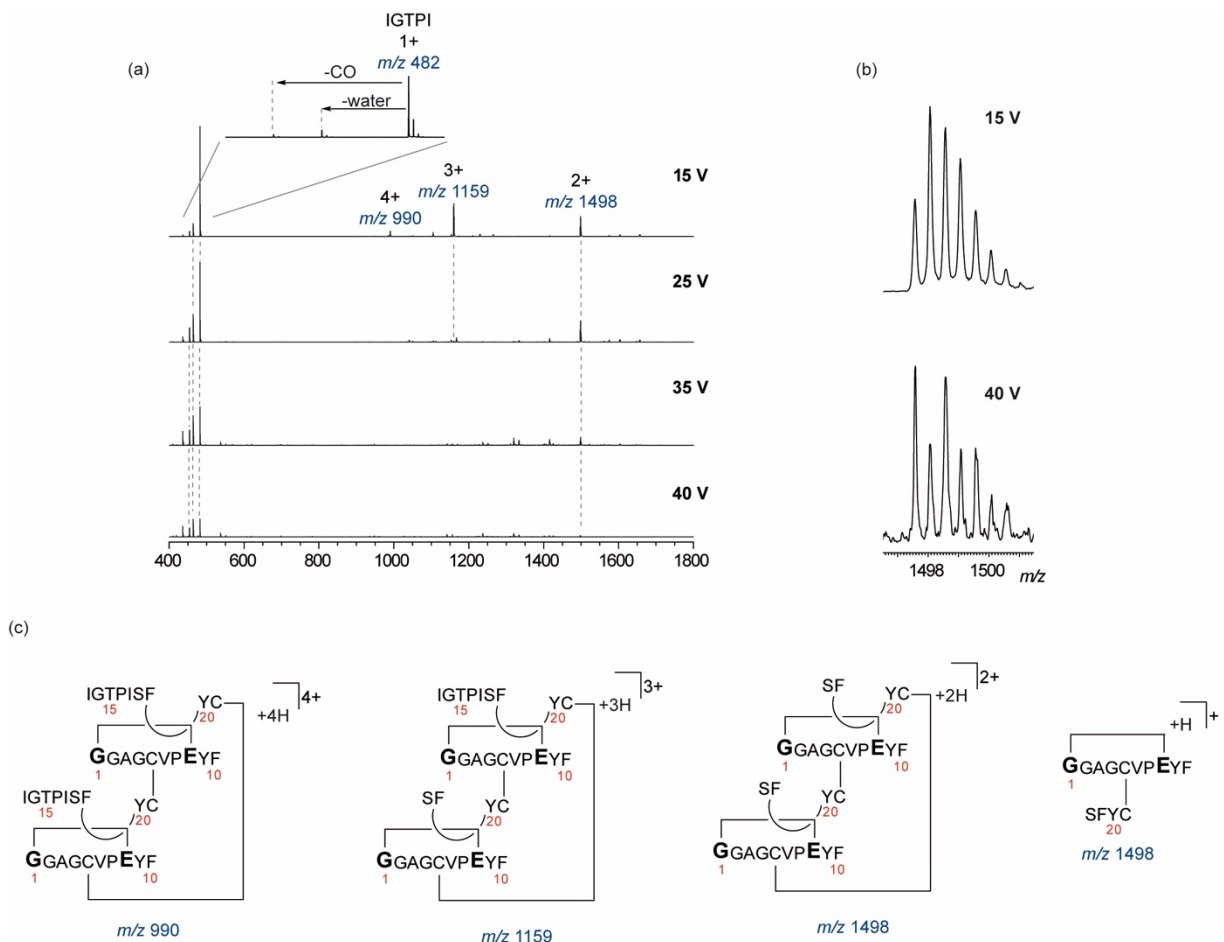
Figure 34. CID experiments with $(2\text{-}2\text{H})_2$. This dimer displays a similar structure to $(7\text{-}3\text{H})_2$ except that it lacks the Cys-13/Cys-13 disulfide bond. (a) CID experiments with mass-selected *m/z* 990 (ion $[\text{M}+4\text{H}]^{4+}$) at different collision voltages. (b) Isotopic patterns of the *m/z* 1498 fragments at collision voltages of 15 and 40 V show the presence of two fragments: (c) A doubly-charged dimeric structure and a singly-charged ring with a Cys-5/Cys-21 connectivity which emerges at higher collision voltages. Thus, it can be assumed that the most abundant dimer from the Cys-21 and Cys-5 building block **2** has two Cys-5/Cys-21 disulfide bonds.

31

6. NMR and MD Simulation Data
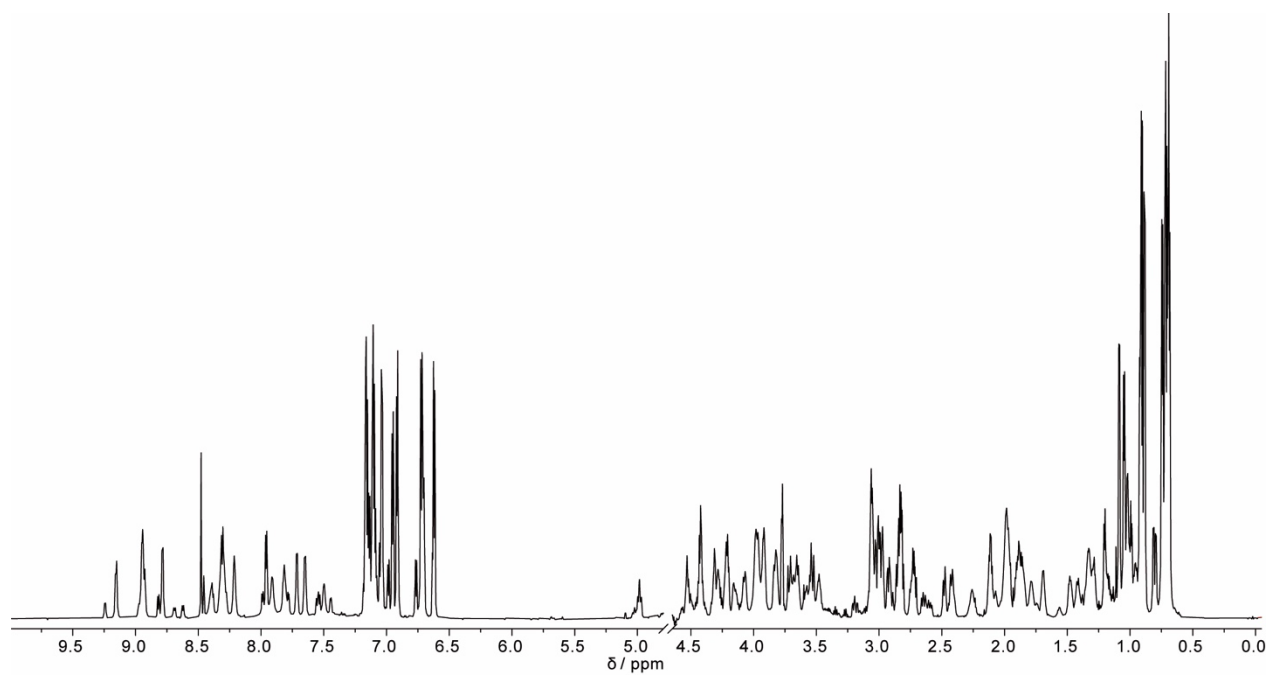
Figure 35. ^{1}H NMR (800 MHz, 295 K) spectrum of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5).
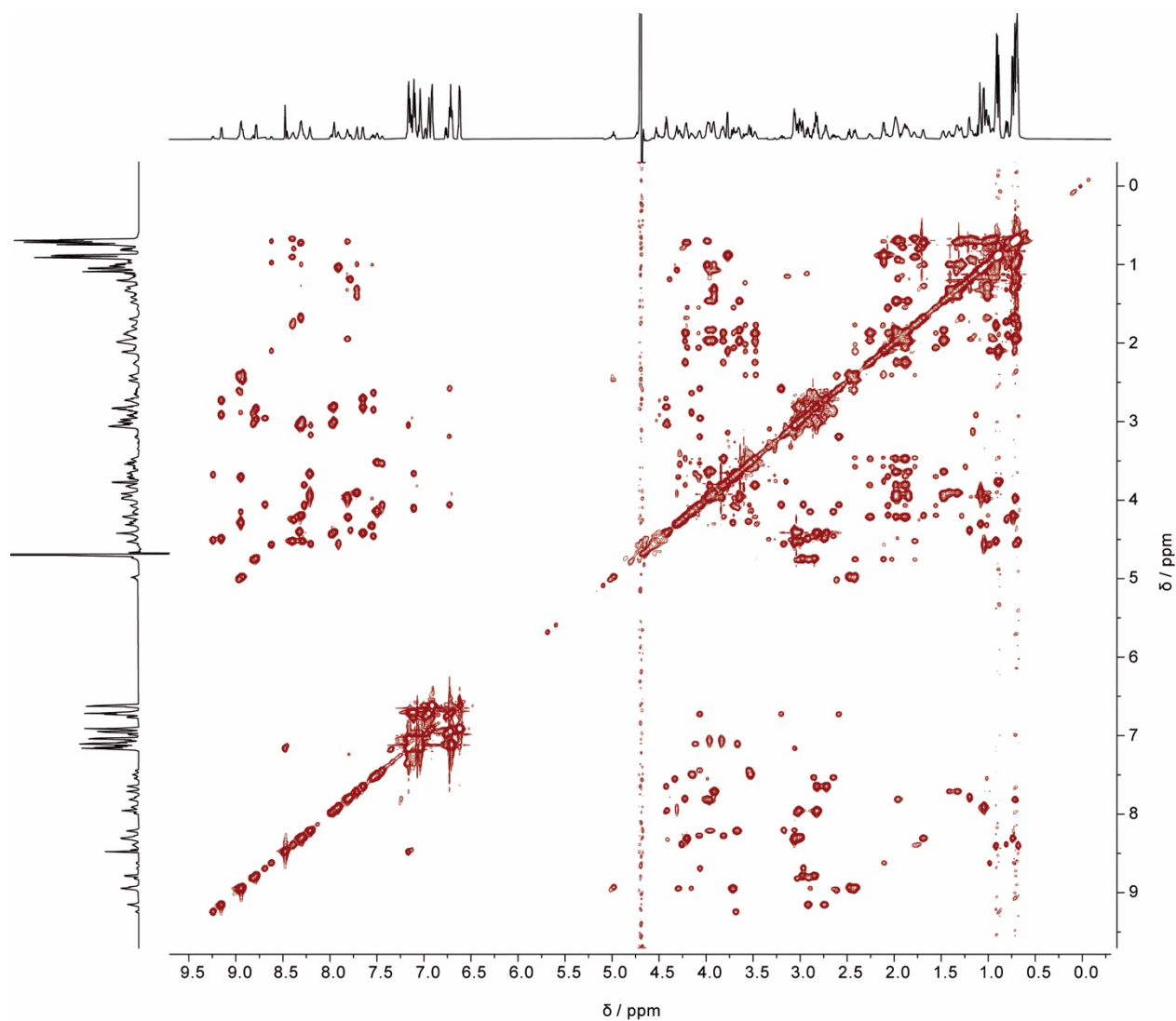
Figure 36. Two-dimensional $^1H,^1H$ TOCSY (800 MHz, 295 K) spectrum of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5). Wide chemical shift dispersion in the fingerprint region indicates a highly structured conformation.

Table 7. Chemical shifts of protons of the major conformer of **3**-2H

residue	proton	shift / ppm	residue	proton	shift / ppm
Gly1	H	'unresolved'	Cys12	H	8.785
Gly1	HA2	'unresolved'	Cys12	HA	4.749
Gly1	HA3	'unresolved'	Cys12	HB2	2.966
Gly2	H	8.947	Cys12	HB3	2.829
Gly2	HA2	4.296	Ile13	H	8.309
Gly2	HA3	3.710	Ile13	HA	4.200
Ala3	H	7.910	Ile13	HB	1.673
Ala3	HA	4.553	Ile13	QG2	0.732
Ala3	QB	1.026	Ile13	HG12	1.261
Gly4	H	7.497	Ile13	HG13	1.006
Gly4	HA2	4.166	Ile13	QD1	0.732
Gly4	HA3	3.515	Gly14	H	8.214
His5	H	8.322	Gly14	HA2	3.946
His5	HA	4.396	Gly14	HA3	3.671
His5	QB	3.051	Thr15	H	7.805
His5	HD1	'exchanged'	Thr15	HA	4.220
His5	HD2	7.159	Thr15	HB	'unresolved'
His5	HE1	8.477	Thr15	QG2	1.174
Val6	H	8.399	Pro16	HA	4.222
Val6	HA	4.514	Pro16	HB2	2.241
Val6	HB	1.771	Pro16	HB3	1.967
Val6	QG1	0.908	Pro16	HG2	1.869
Val6	QG2	0.673	Pro16	HG3	1.869
Pro7	HA	3.974	Pro16	HD2	3.817
Pro7	HB2	1.967	Pro16	HD3	3.475
Pro7	HB3	1.829	Ile17	H	7.814
Pro7	QG	1.457	Ile17	HA	3.967
Pro7	HD2	3.925	Ile17	HB	1.947
Pro7	HD3	3.646	Ile17	QG2	0.7125
Glu8	H	7.713	Ile17	HG12	1.125
Glu8	HA	3.906	Ile17	HG13	'unresolved'
Glu8	QB	0.987	Ile17	QD1	0.7125
Glu8	HG2	1.398	Ser18	H	'unresolved'
Glu8	HG3	1.320	Ser18	HA	'unresolved'
Tyr9	H	7.648	Ser18	HB2	'unresolved'
Tyr9	HA	4.416	Ser18	HB3	'unresolved'
Tyr9	HB2	2.809	Phe19	H	8.931
Tyr9	HB3	2.711	Phe19	HA	4.984
Tyr9	HD1	6.914	Phe19	HB2	2.456
Tyr9	HD2	6.914	Phe19	HB3	2.398
Tyr9	HE1	6.618	Phe19	HD1	6.704
Tyr9	HE2	6.618	Phe19	HD2	6.704
Tyr9	HH	'exchanged'	Phe19	HE1	7.107
Phe10	H	7.957	Phe19	HE2	7.107
Phe10	HA	4.416	Phe19	HZ	7.161
Phe10	HB2	3.005	Tyr20	H	9.154
Phe10	HB3	2.809	Tyr20	HA	4.494
Phe10	HD1	7.036	Tyr20	HB2	2.907
Phe10	HD2	7.036	Tyr20	HB3	2.731
Phe10	HE1	7.152	Tyr20	HD1	6.948
Phe10	HE2	7.152	Tyr20	HD2	6.948
Phe10	HZ	7.090	Tyr20	HE1	6.719
Val11	H	'exchanged'	Tyr20	HE2	6.719
Val11	HA	3.770	Tyr20	HH	'exchanged'
Val11	HB	2.104	Cys21	H	8.293
Val11	QG1	0.889	Cys21	HA	4.514
Val11	QG2	0.889	Cys21	HB2	3.044
			Cys21	HB3	2.985
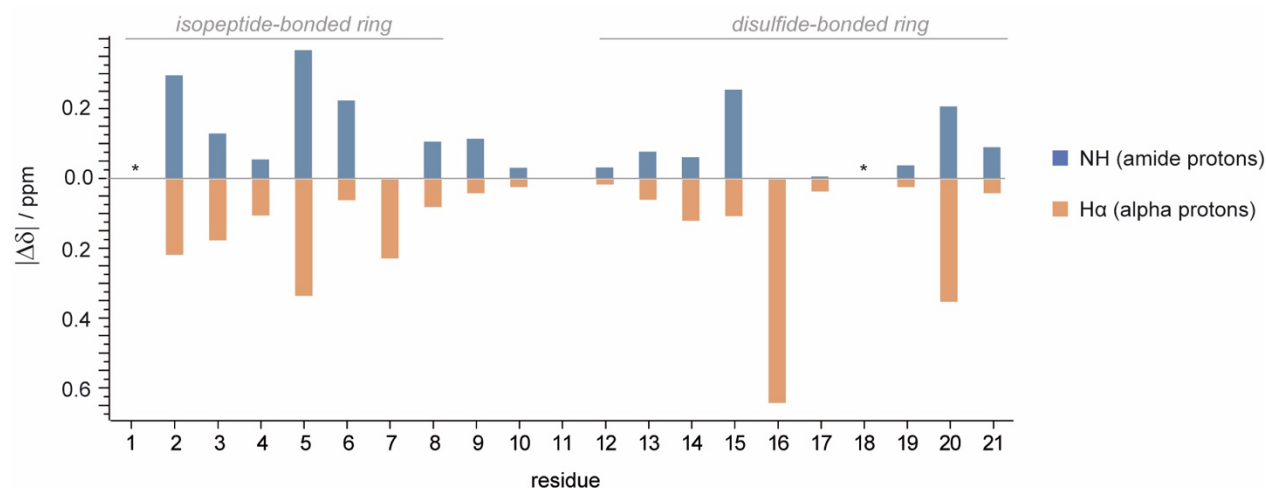
Figure 37. Plot of chemical shift differences between the amide protons of the major and minor conformer of [2]catenane **3**-2H. Protons which could not be unambiguously assigned are marked with an asterisk. Only small shift differences are observed for residues Phe-9–Val-11, which are not members of the peptide macrocycles.
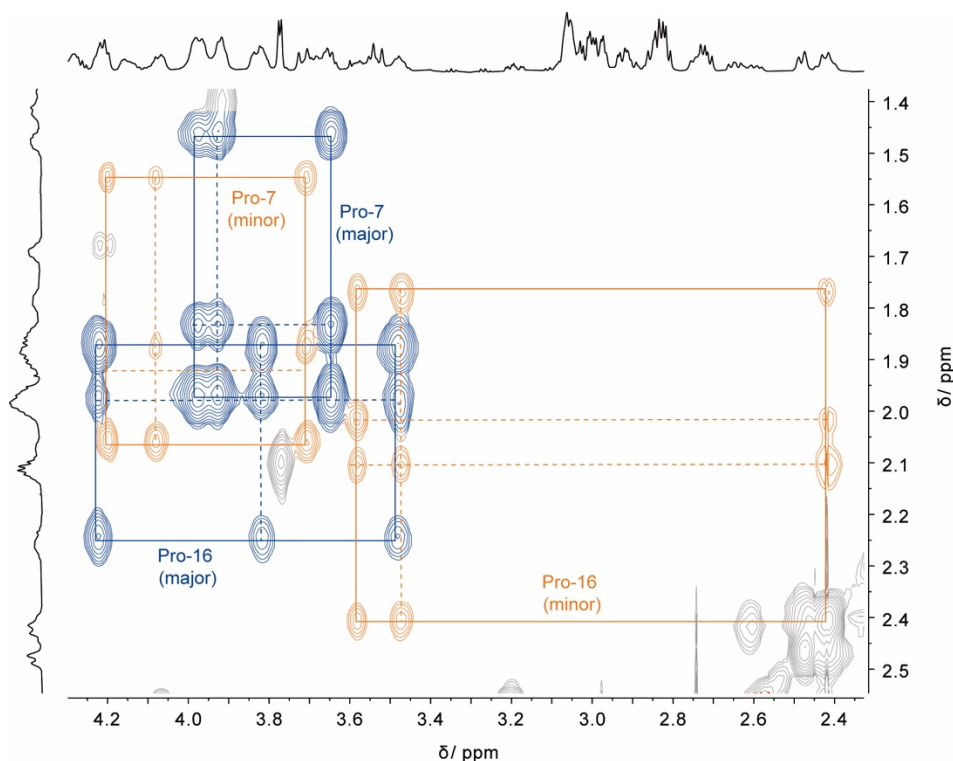
Figure 38. Partial ^{1}H,^{1}H TOCSY (800 MHz, 295 K) spectrum of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5). The off-diagonal cross peaks related to the TOCSY-pattern of the Pro-7 and Pro-16 residue are colored in blue (major conformer) and orange (minor conformer).
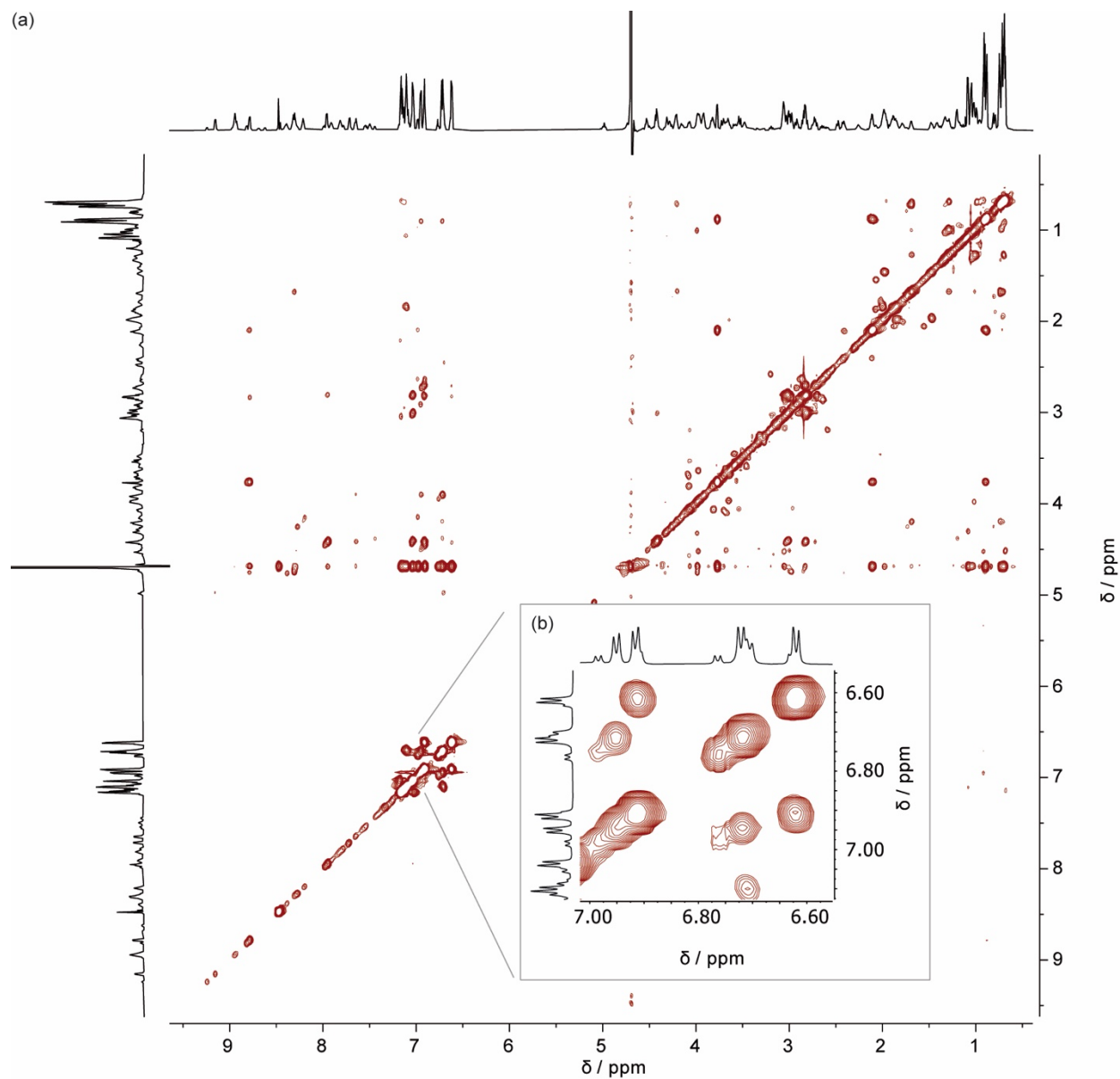
Figure 39. (a) Two-dimensional $^{1}H,^{1}H$ EXSY spectrum (800 MHz, 295 K, t_m = 1.5 s) of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5). (b) Partial spectrum showing the Tyr-9 and Tyr-20 protons. The absences of off-diagonal cross peaks between the signal sets of the major and minor conformer indicates slow exchange on the NMR timescale.
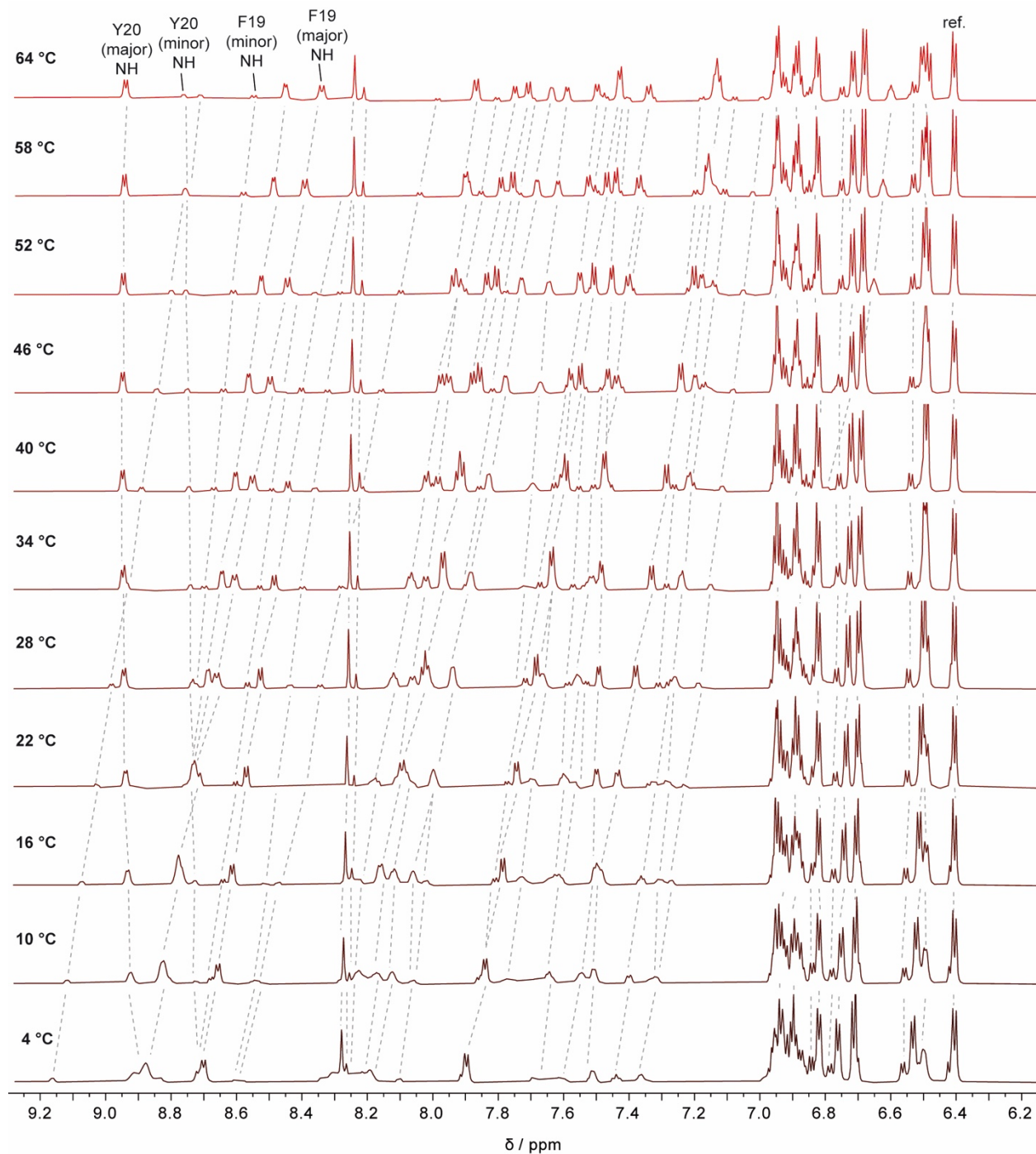
Figure 40. Variable temperature NMR experiment: partial ^{1}H NMR (800 MHz) spectra of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5) at different temperatures. Reverse up- and downfield shifts, as exemplified on the selected amide protons of the lock residues Phe-19 and Tyr-20 of both conformers, indicate changes of local chemical/magnetic enviroments during heating. All spectra are internally referenced to the aromatic signals of residue Tyr-9 (Hε = 6.40 ppm).
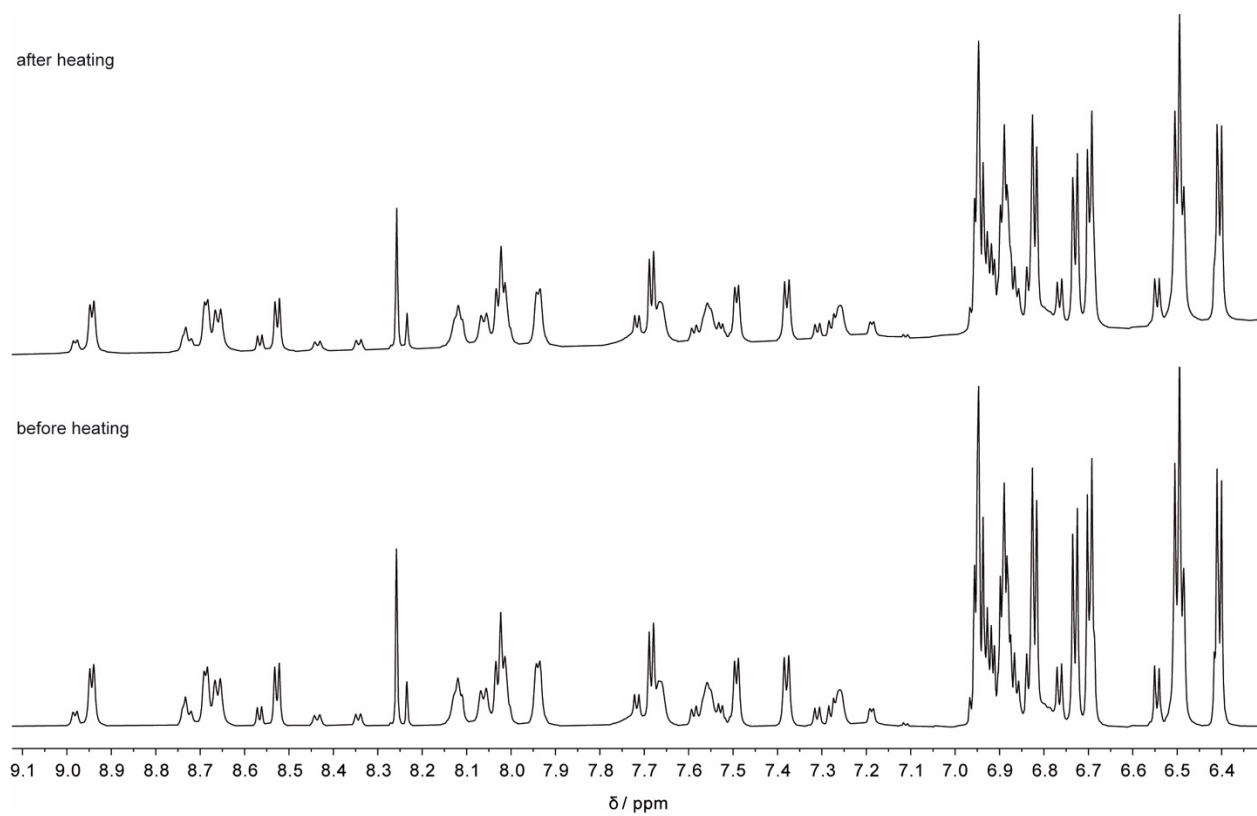
Figure 41. Partial ^{1}H NMR (800 MHz, 295 K) spectra of [2]catenane **3**-2H (1.03 mM) in H$_2$O/D$_2$O (95/5) before and after heating to 64 °C.
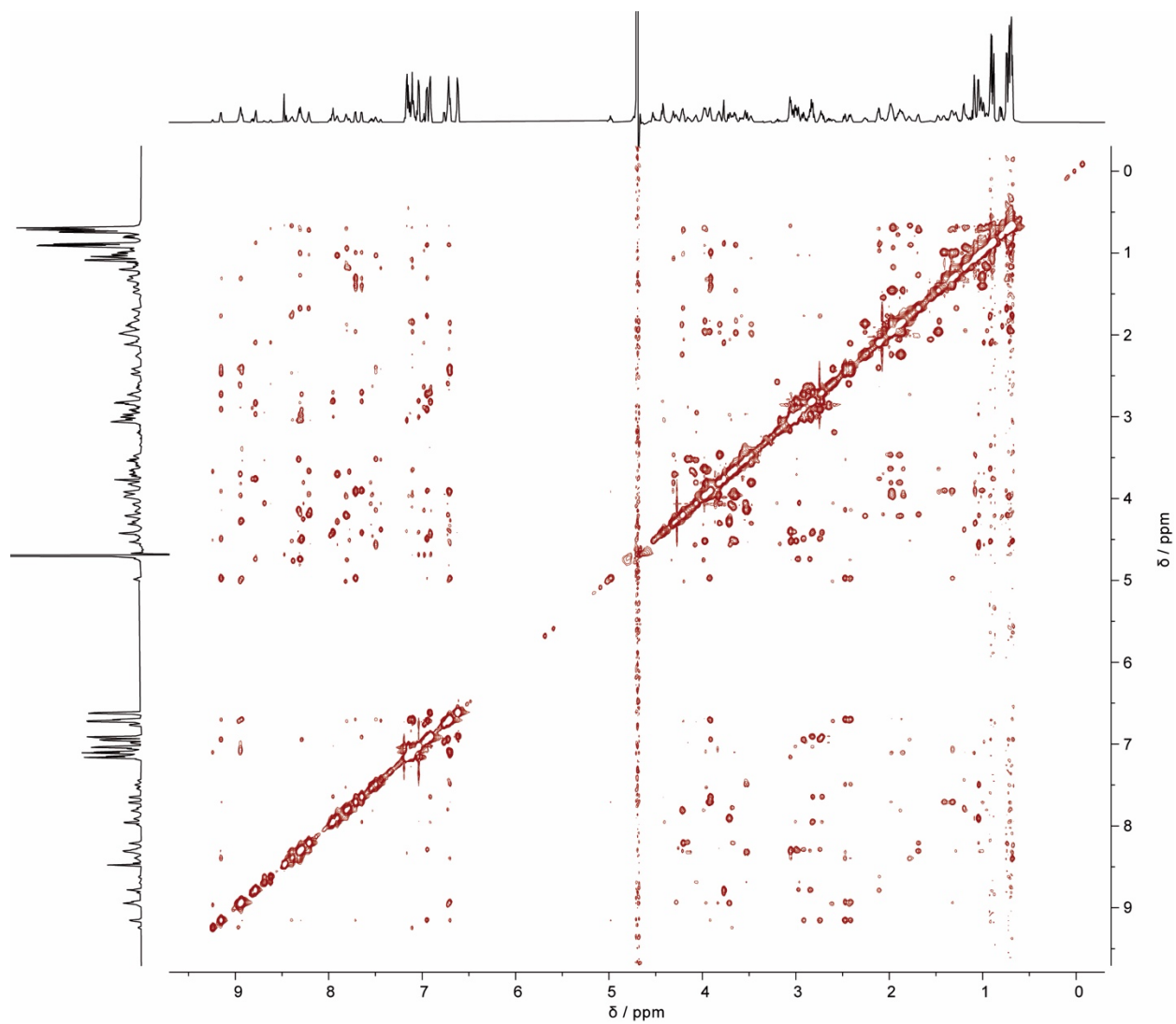
Figure 42. ^{1}H,^{1}H NOESY (800 MHz, 295 K, t_m = 300 ms) spectrum of [2]catenane **3**-2H (1.03 mM) in H_2O/D_2O (95/5).
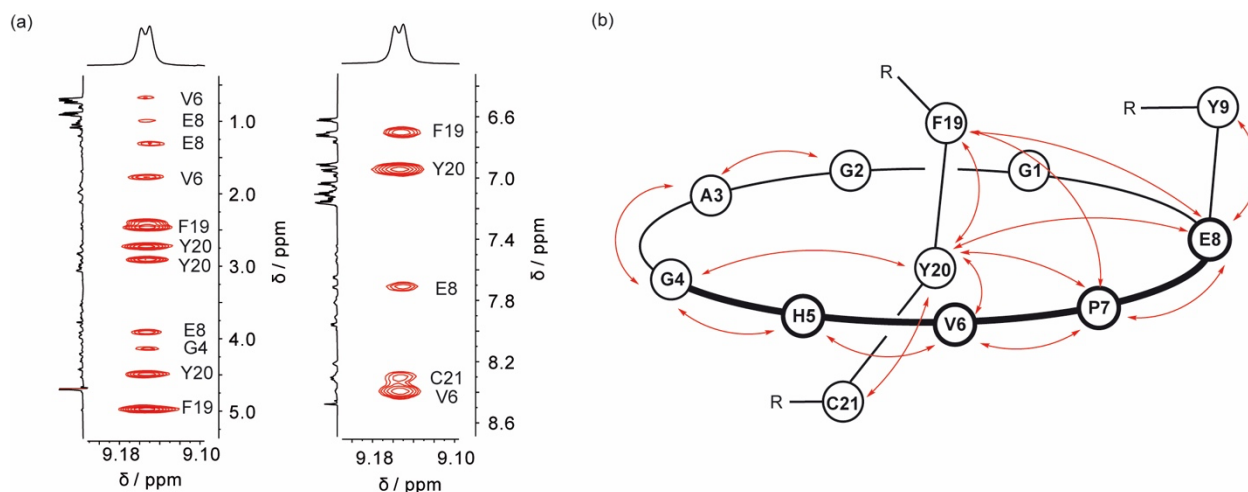
Figure 43. Intra- and interchain NOE correlations (see Figure 42 for full spectrum) observed for the isopeptide-bonded macrocycle and lock residues Phe-19 and Tyr-20: (a) As an example, the amide proton (NH) of Tyr-20 shows several NOE interactions with adjacent residues in both rings. (b) Schematic illustration showing selected interresidue NOE correlations.
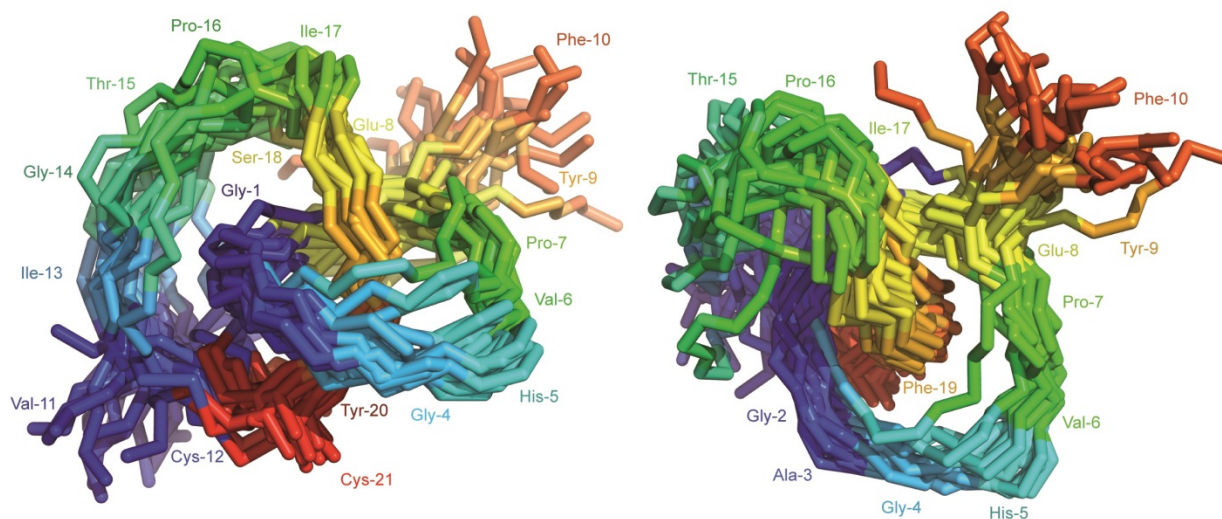
Figure 44. Conformational ensemble showing the superposition of the top 20 energy-minimized conformations of [2]catenane **3**-2H in two different views. The backbone of all residues and the side chains of Cys-12, Cys-21 and Glu-8 are shown as sticks. The models are structurally aligned to minimize the root mean square deviation (RMSD) of the whole peptide backbone.
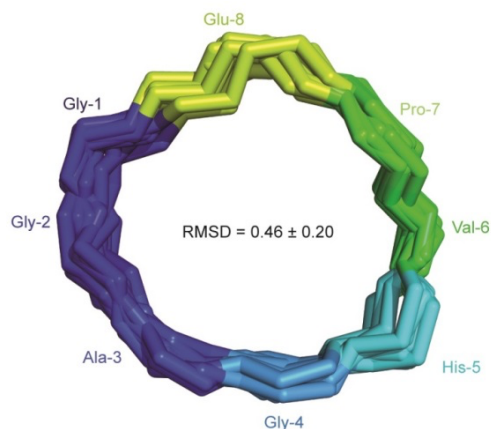
Figure 45. Superposition of the top 20 energy-minimized conformations of the isopeptide bonded ring in the NMR ensmeble of [2]catenane **3**-2H. The backbone of residues 1–8 and the side chain of Glu-8 are shown as sticks and oxygen and hydogen atoms are removed for the sake of clarity. The models are structurally aligned to minimize the root mean square deviation (RMSD) of the ring residue backbone (residue Gly-1–Glu-8).
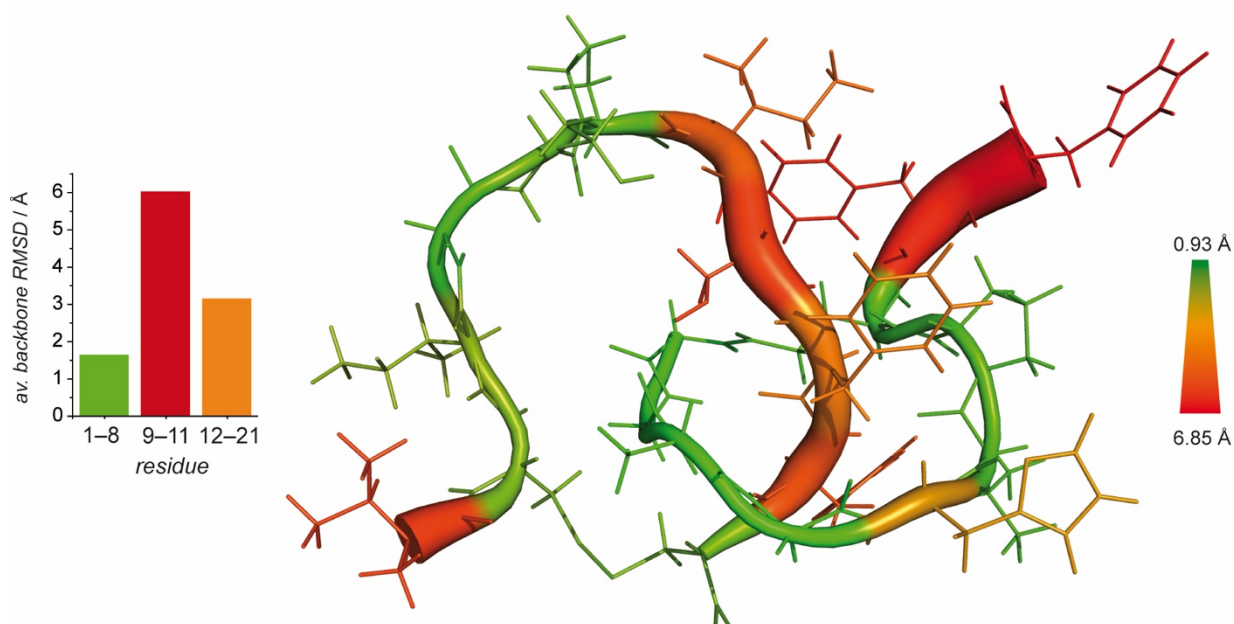
Figure 46. MD simulation-derived (1 μs, 298 K) *B*-factor mapping of residues in [2]catenane **3**-2H plotted on the lowest-energy conformer (structure 1 from the NMR ensemble). The plot is based on the RMSD values of the peptides' backbone atoms. The RMSD value range is shown as cartoon-like backbone structure as well as color gradient ranging from 0.93 (green) to 6.85 (red) Å. The residues are accordingly colored and shown as lines. Note that the plot describes the individual confomational flexibility of each ring without taking co-conformational motion into account.

7. Mass Spectra

7.1. Lasso Peptides
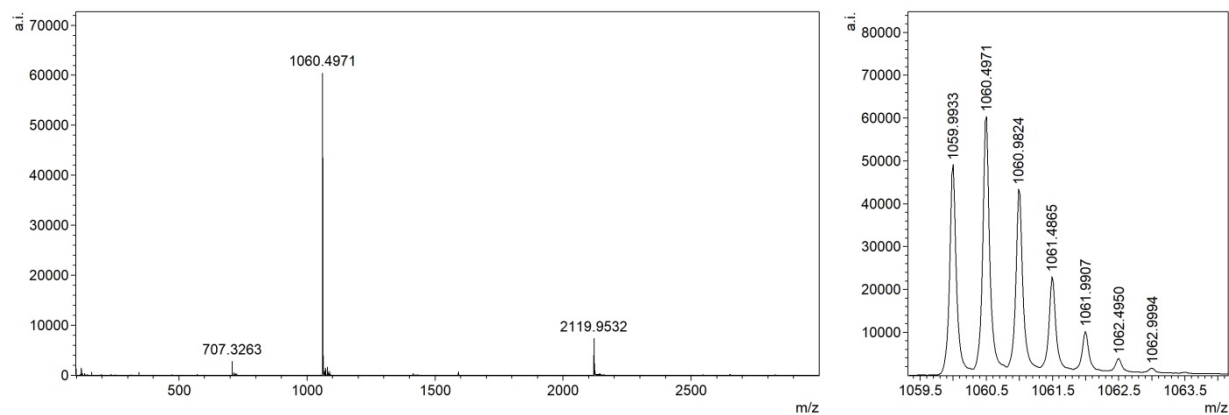
Figure 47. ESI mass spectrum of peptide **L2** and isotopic distribution of the [M+2H]$^{2+}$ ion.
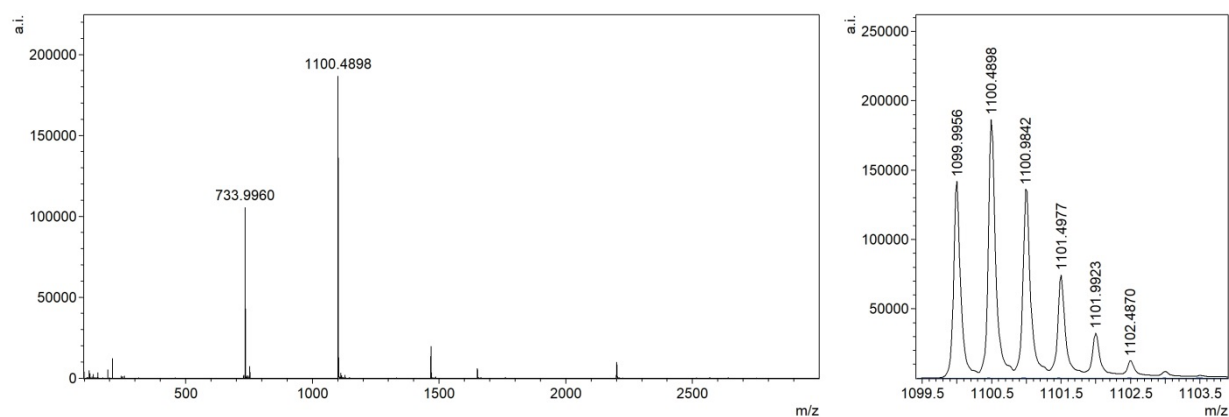
Figure 48. ESI mass spectrum of peptide **L3** and isotopic distribution of the [M+2H]$^{2+}$ ion.
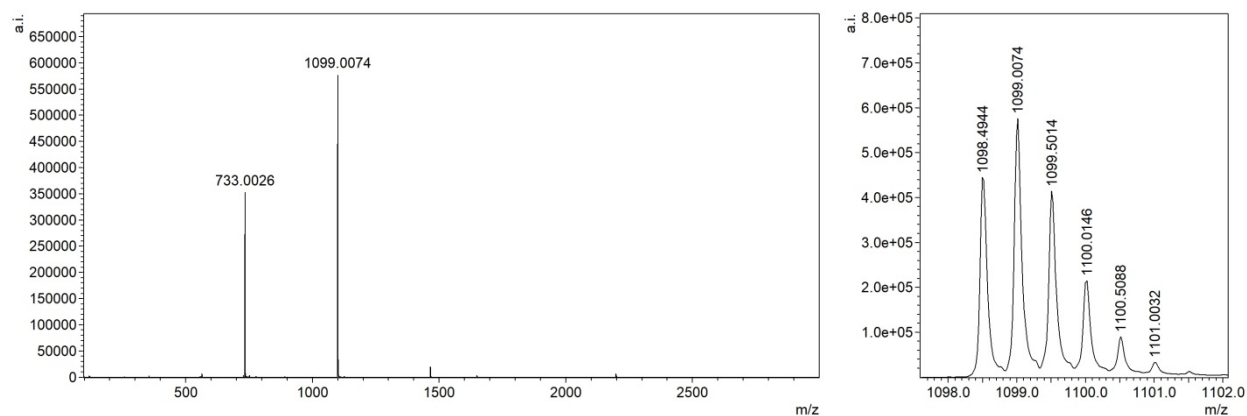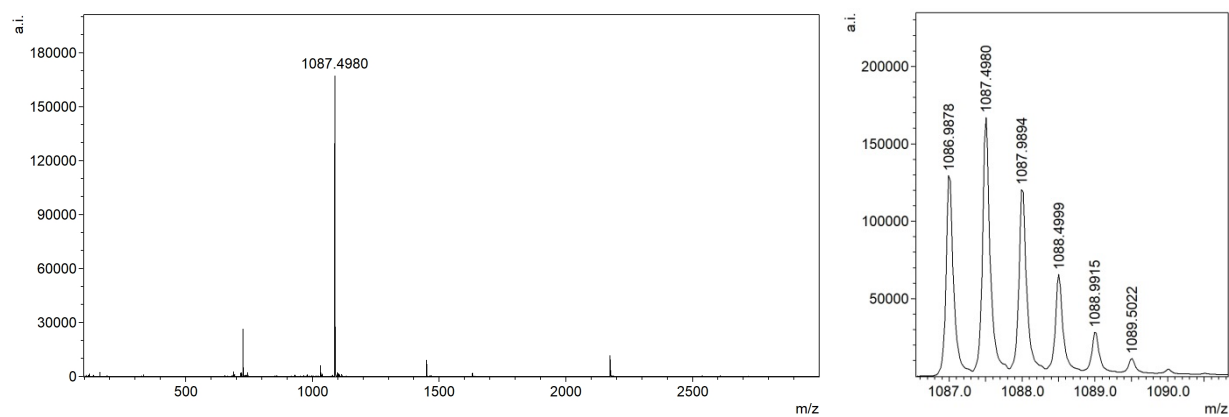

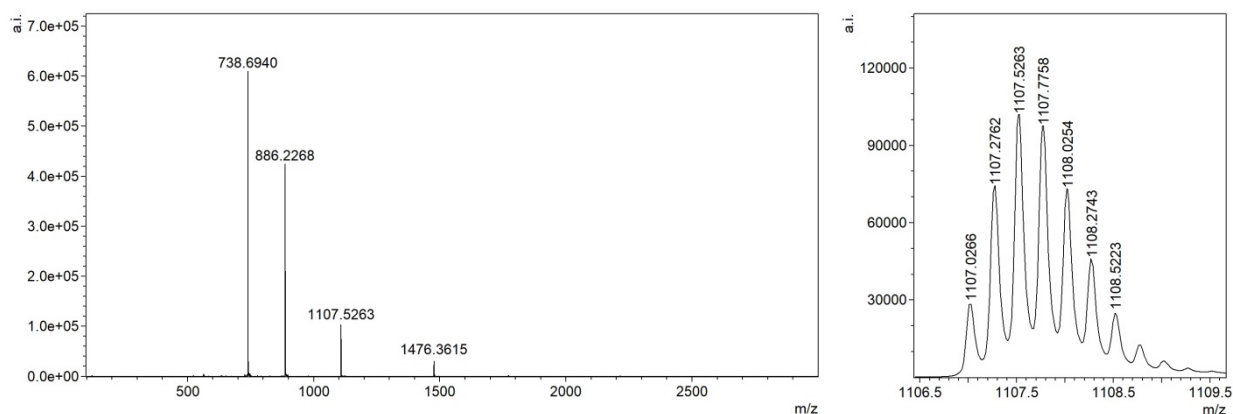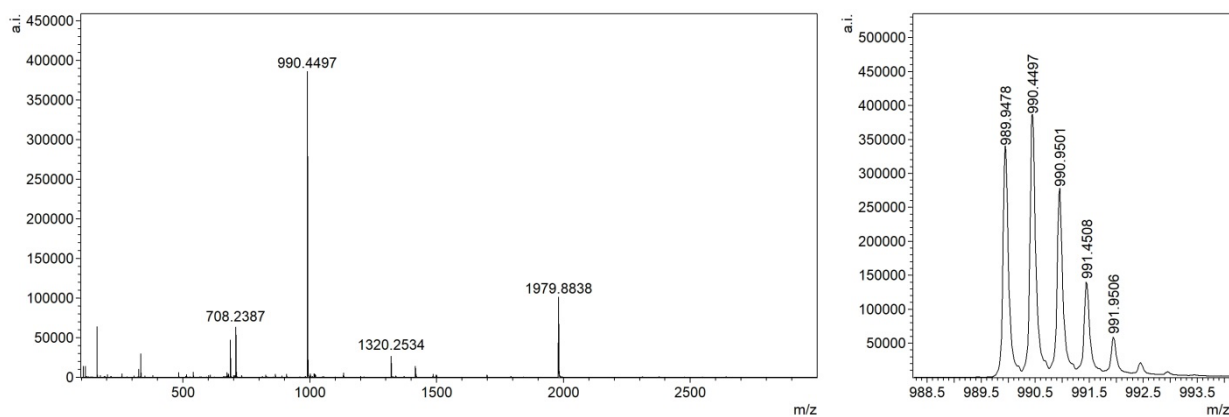**Figure 49.** ESI mass spectrum of peptide **L4** and isotopic distribution of the [M+2H]$^{2+}$ ion.
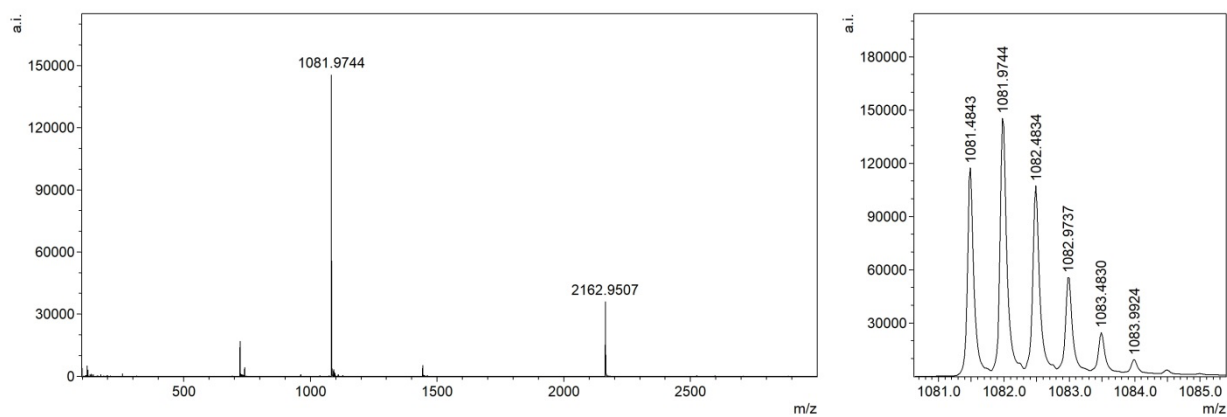
Figure 50. ESI mass spectrum of peptide **L5** and isotopic distribution of the [M+2H]$^{2+}$ ion.
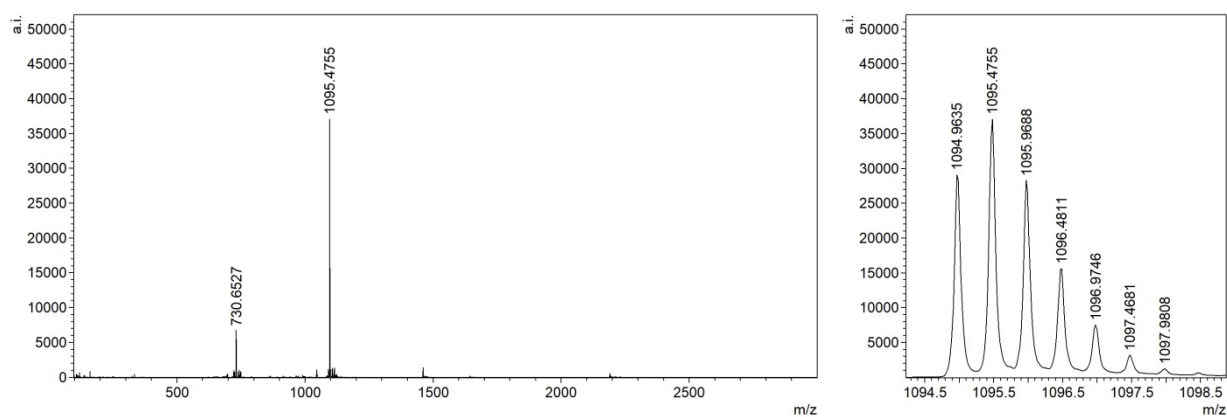
Figure 51. ESI mass spectrum of peptide **L6** and isotopic distribution of the [M+2H]$^{2+}$ ion.
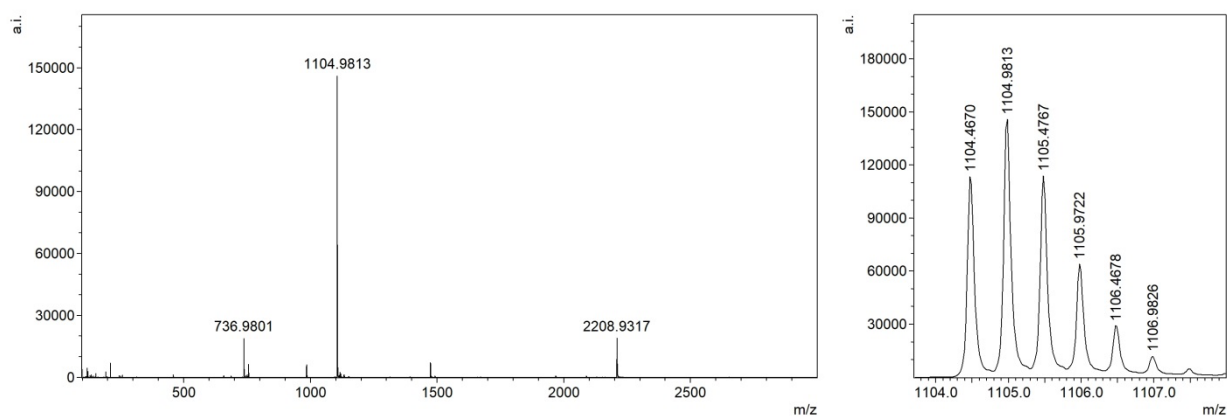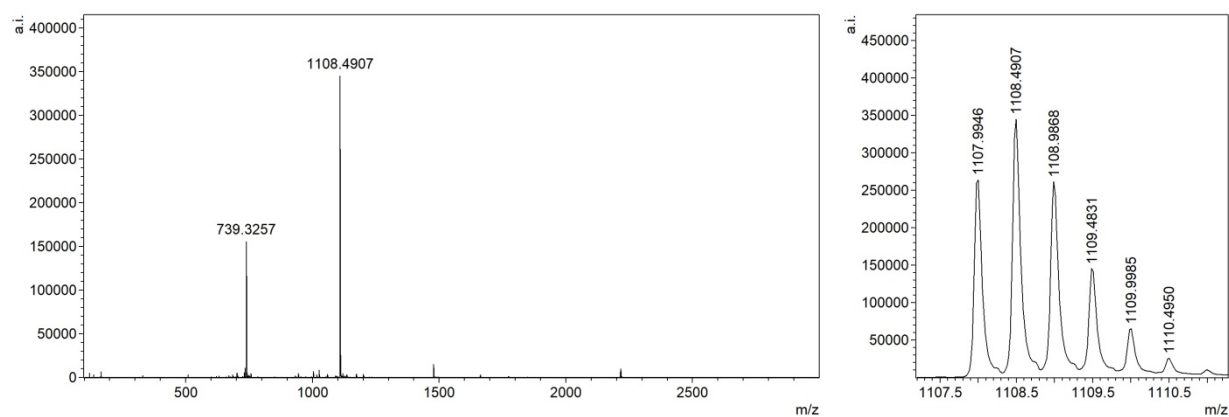

Figure 52. ESI mass spectrum of peptide **L7** and isotopic distribution of the [M+2H]$^{2+}$ ion.
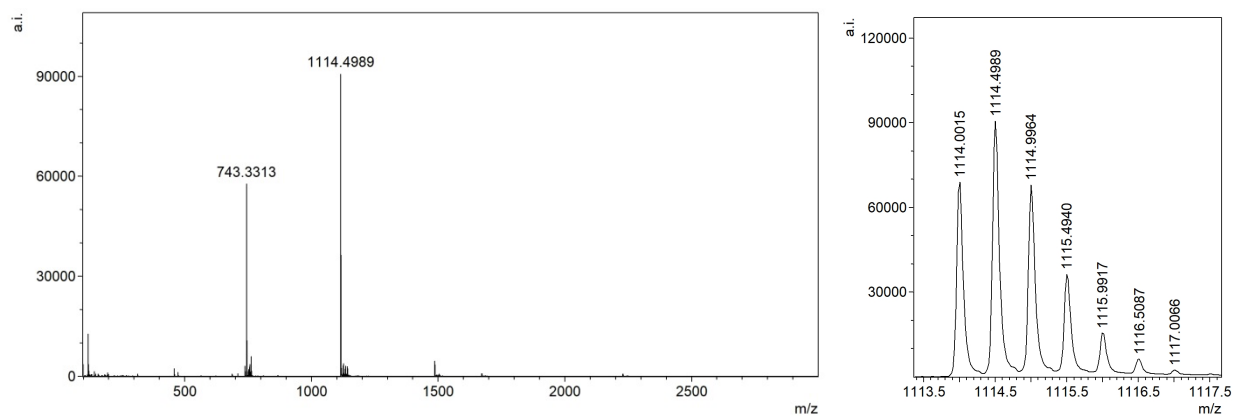
Figure 53. ESI mass spectrum of peptide **L8** and isotopic distribution of the $[M+2H]^{2+}$ ion.


Figure 54. ESI mass spectrum of peptide **L3 P16A** variant and isotopic distribution of the $[M+2H]^{2+}$ ion.

7.2. Mechanically Interlocked Peptides
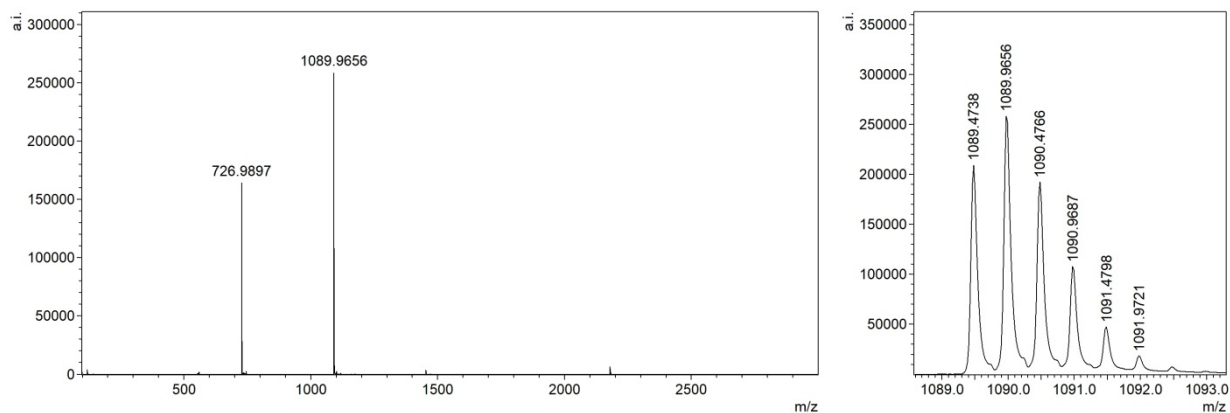
Figure 55. ESI mass spectrum of MIP **5**-2H and isotopic distribution of the [M+2H]$^{2+}$ ion.
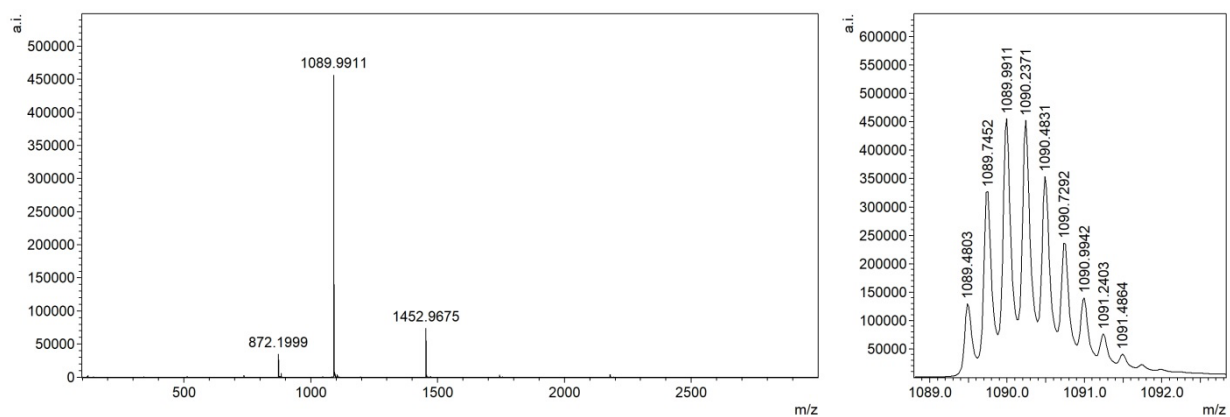
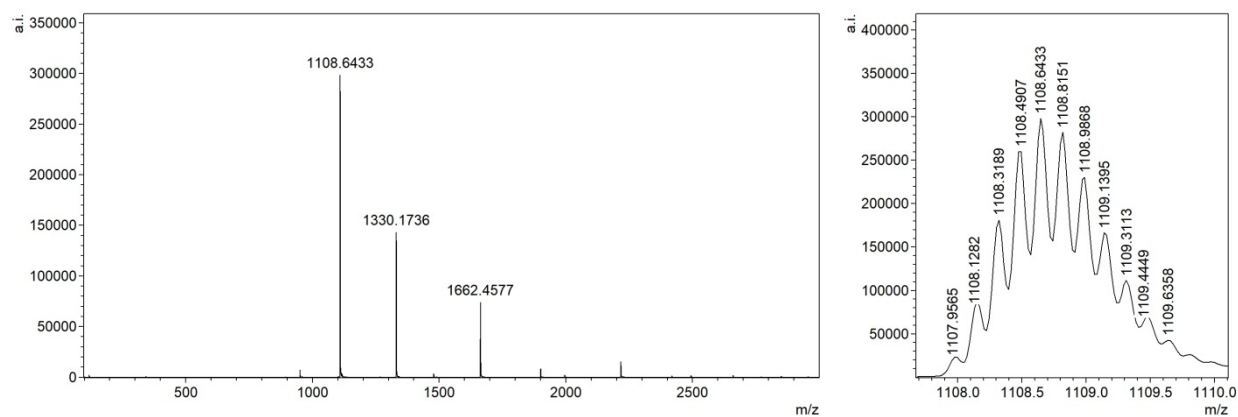**Figure 56.** ESI mass spectrum of MIP (**5**-2H)$_2$ and isotopic distribution of the [M+4H]$^{4+}$ ion.

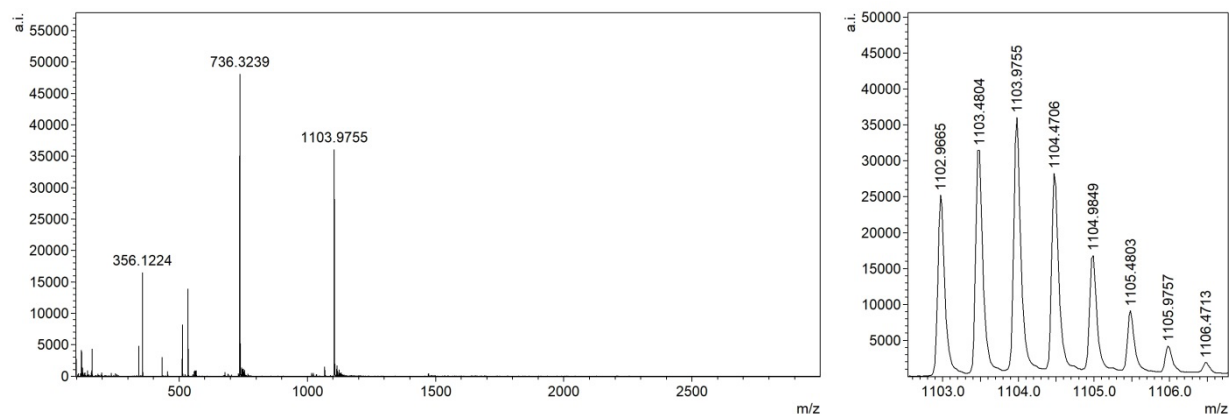**Figure 57.** ESI mass spectrum of MIP (**4**-H)$_2$ and isotopic distribution of the [M+4H]$^{4+}$ ion.
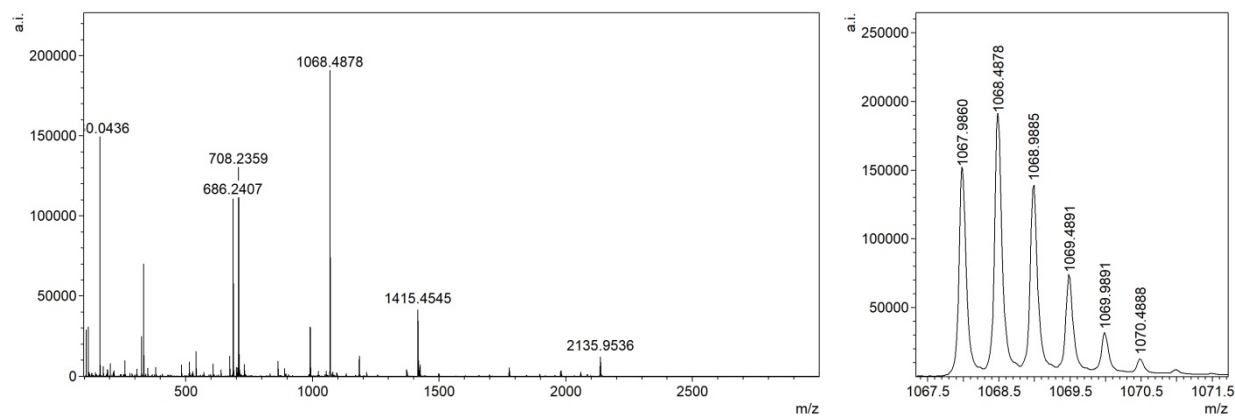
Figure 58. ESI mass spectrum of MIP **2**-2H and isotopic distribution of the [M+2H]$^{2+}$ ion.

Figure 59. ESI mass spectrum of MIP **2**-2H (-ValGly) and isotopic distribution of the [M+2H]$^{2+}$ ion.

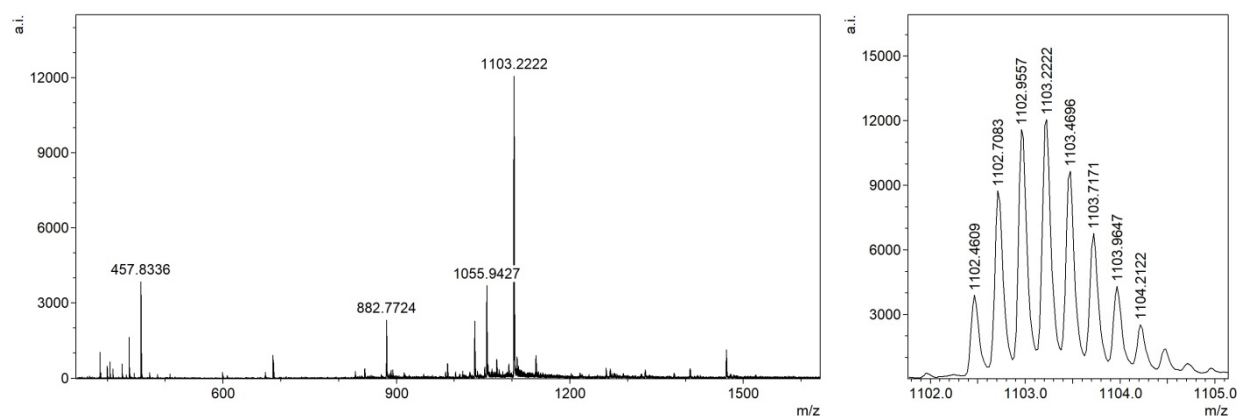**Figure 60.** ESI mass spectrum of MIP **3**-2H and isotopic distribution of the [M+2H]$^{2+}$ ion.
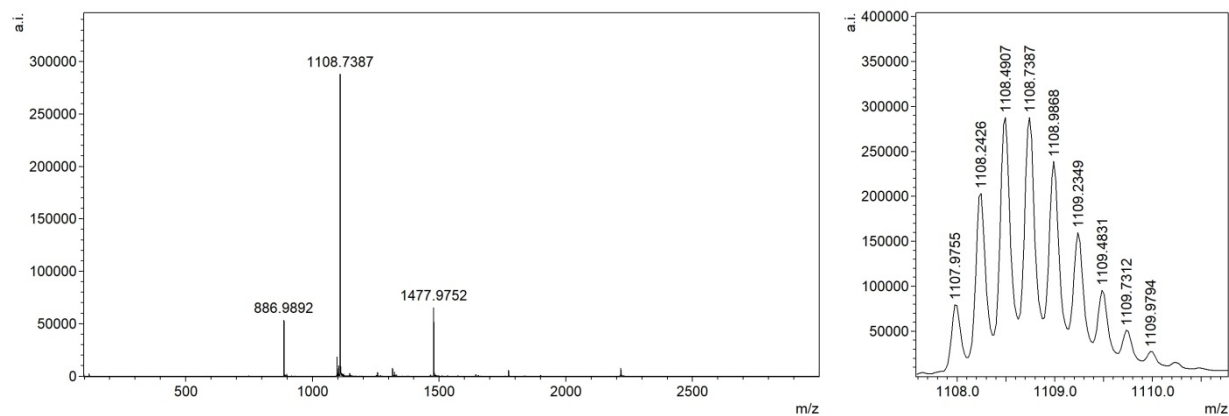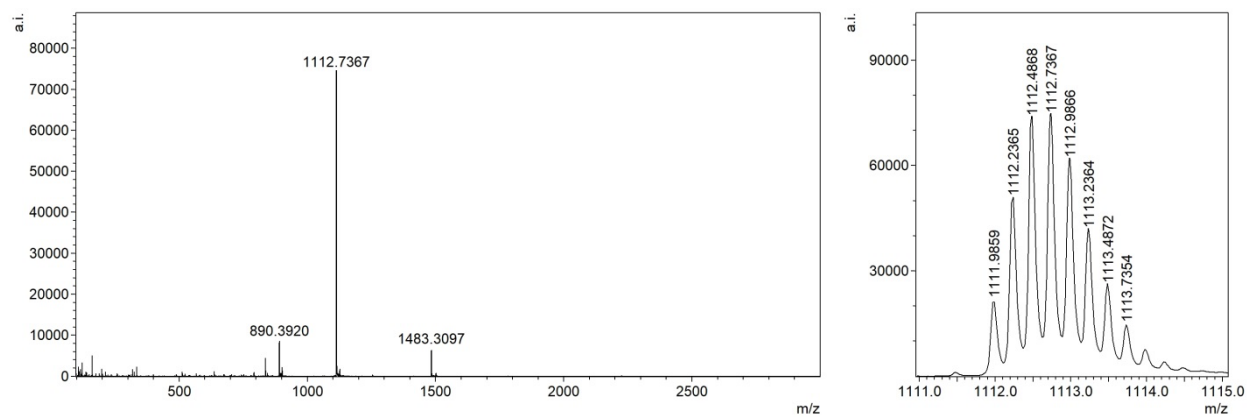
Figure 61. ESI mass spectrum of MIP (3-2H)$_2$ and isotopic distribution of the [M+4H]$^{4+}$ ion.

Figure 62. ESI mass spectrum of MIP (3-2H)$_3$ and isotopic distribution of the [M+6H]$^{6+}$ ion.
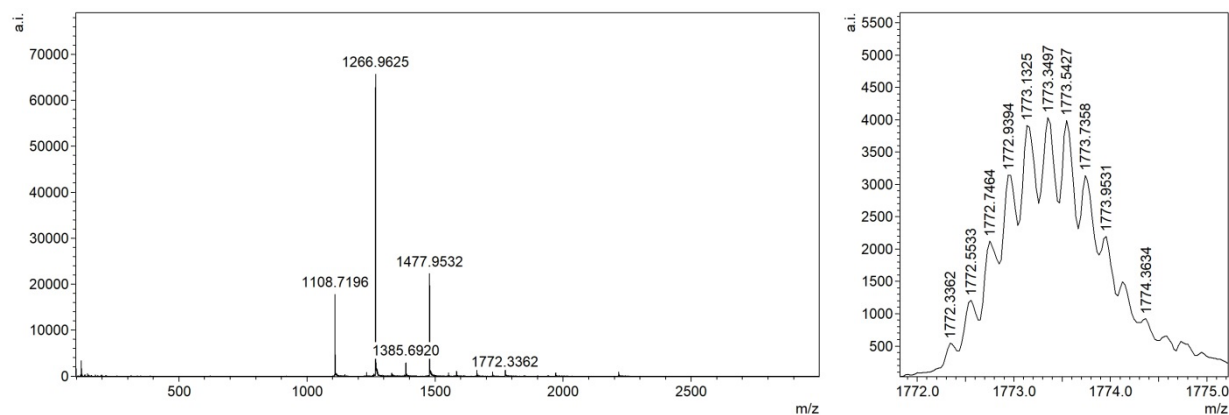
Figure 63. ESI mass spectrum of MIP (3-2H)$_4$ and isotopic distribution of the [M+5H]$^{5+}$ ion.

Figure 64. ESI mass spectrum of MIP **6**-2H and isotopic distribution of the $[M+2H]^{2+}$ ion.
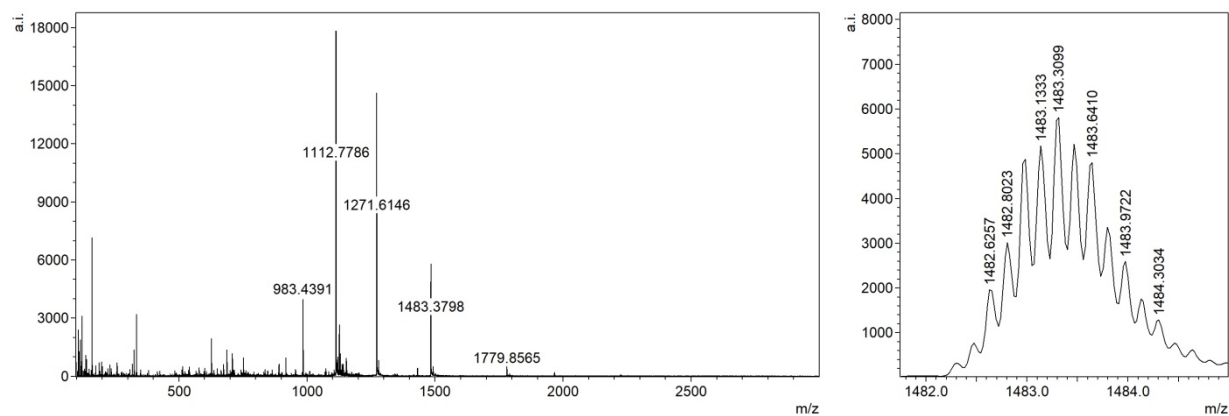

Figure 65. ESI mass spectrum of MIP (**6**-3H)$_2$ and isotopic distribution of the $[M+4H]^{4+}$ ion.

Figure 66. ESI mass spectrum of MIP (**7**-3H)$_2$ and isotopic distribution of the $[M+4H]^{4+}$ ion.

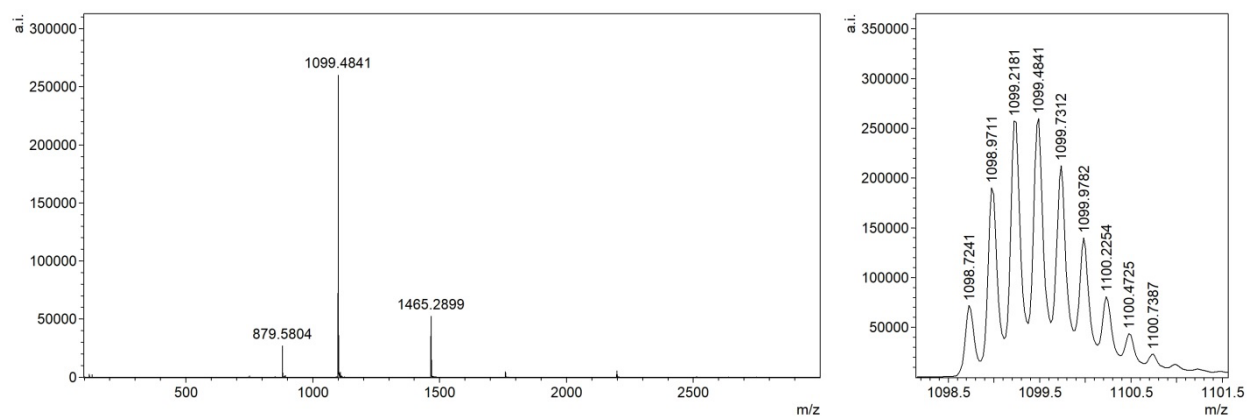**Figure 67.** ESI mass spectrum of MIP (**7**-3H)$_4$ and isotopic distribution of the [M+6H]$^{6+}$ ion.

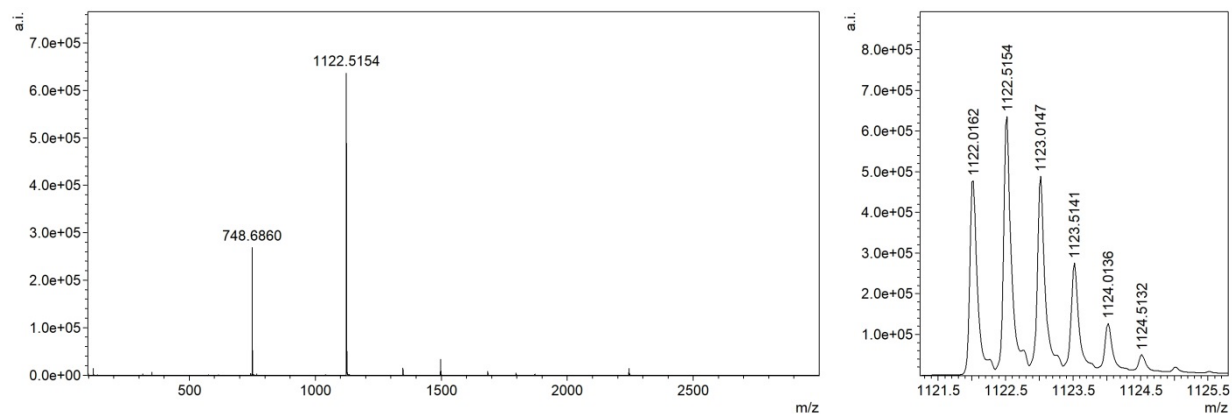**Figure 68.** ESI mass spectrum of MIP **3•5**-4H and isotopic distribution of the [M+4H]$^{4+}$ ion.

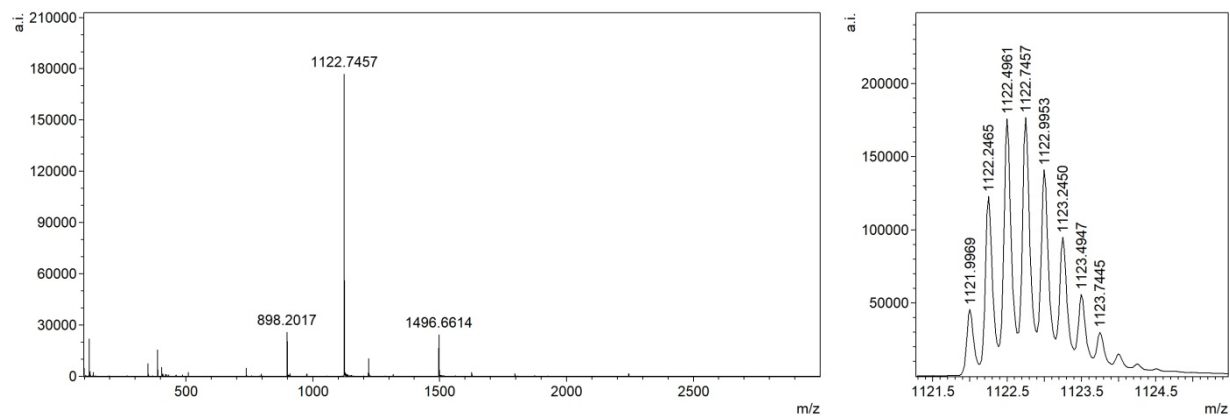**Figure 69.** ESI mass spectrum of MIP **8**-2H and isotopic distribution of the [M+2H]$^{2+}$ ion.

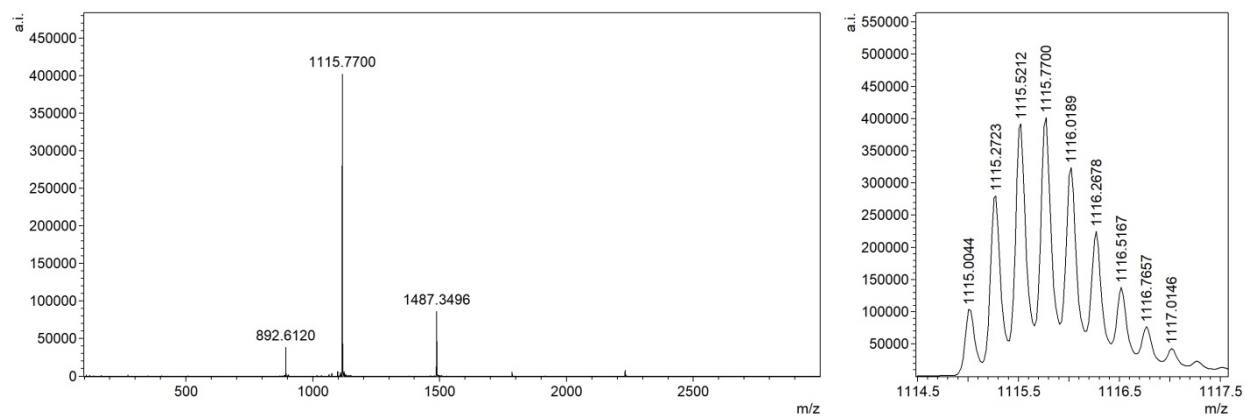**Figure 70.** ESI mass spectrum of MIP (**8**-2H)$_2$ and isotopic distribution of the [M+4H]$^{4+}$ ion.

Figure 71. ESI mass spectrum of MIP **3•8**-4H and isotopic distribution of the [M+4H]$^{4+}$ ion.
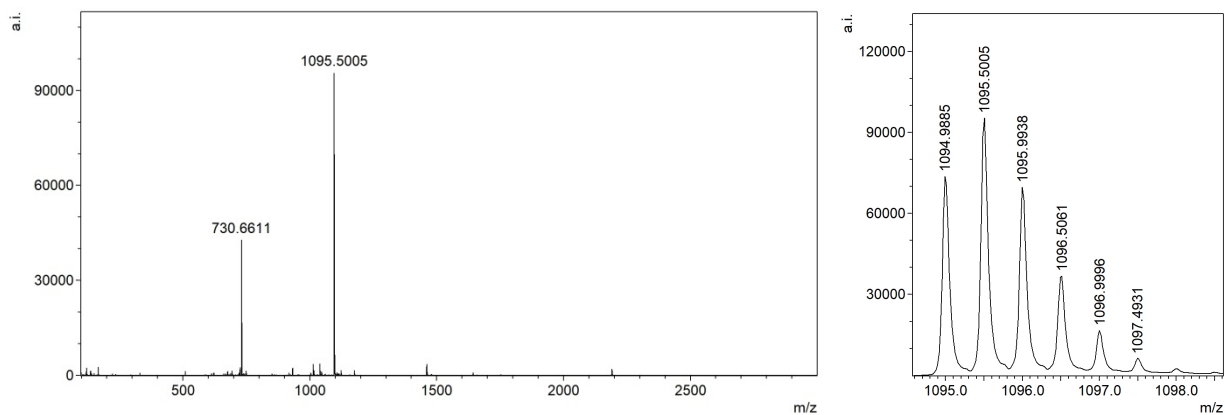

Figure 72. ESI mass spectrum of MIP **3**-2H **P16A** and isotopic distribution of the [M+2H]$^{2+}$ ion.

8. Supporting Information References

1. C. D. Allen, A. J. Link, *J. Am. Chem. Soc.* **2016**, *138*, 14214–14217.

2. P. Güntert, L. Buchner, *J. Biomol. NMR* **2015**, 62, 453–471.

3. C. A. Spronk, A. M. J. J. Bonvin, P. K. Radha, G. Melacini, R. Boelens, R. Kaptein, *Structure* **1999**, 7, 1483–1492.

4. C. A. Spronk, J. P. Linge, C. W. Hilbers and G. W. Vuister, *J. Biomol. NMR* **2002**, 22, 281–289.

5. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781–1802.

6. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, 14, 33–38.

7. P.-K. Lai, Y. N. Kaznessis, *J. Chem. Theory Comput.* **2017**, 13, 3413–3423.

8   C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*,1652–1671.

9. V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

10. F. Neese, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.

11. C. Bleiholder, S. Suhai, A. G. Harrison, B. Paizs, *J. Am. Soc. Mass Spectrom.* **2011**, 22, 1032–1039.